

What is public understanding for?

Britain's second annual 'science week' provokes questions about the aims of the public understanding of science — and about the defects of the educational system, still unreformed.

THE idea that there should be a general understanding of science has swiftly become politically correct. In Britain, the modern phase of the movement began a decade ago, when the Royal Society commissioned Sir Walter Bodmer to produce a working paper on the subject and then appointed him the chairman of a Committee on the Public Understanding of Science, called COPUS. (The present chairman is Professor Lewis Wolpert.) Left to its own resources, COPUS would no doubt have continued with the portfolio of worthy activities with which it began — media internships for scientists, the lobbying of newspaper editors and related moguls and the sponsorship of public lectures. All that changed with the arrival, three years ago, of Mr William Waldegrave as the minister of science and technology. (He is now the British agriculture minister.)

The COPUS budget was enlarged and the Office of Science and Technology (OST) began spending money on its own account as well. Others, notably the Wellcome Trust, followed suit. (The trust says it will spend even more now that its income is likely to be doubled.) Public and charitable funds labelled 'public understanding' now probably exceed £10 million a year in Britain. One welcome side-effect has been to give the British Association for the Advancement of Science a sense of purpose (and a more solid public subvention) that it has lacked for too long. Now OST has set up a committee of its own (which shows that it must be serious) to coordinate policy in the field; the chairman is Sir Arnold Wolfendale, Astronomer Royal for England until recently. And 'British Science Week', organized in something of a hurry last year, seems now to be institutionalized. The technical organizations asked to help have responded at least dutifully, sometimes enthusiastically. The varied events of the past few days will have demonstrated both the extent of Britain's ambitions in this field and, often, the inexpertness of their execution (see page 296).

Intention

Nobody can object to the broadening and deepening of the public understanding of science, but what is intended? Without an understanding of what all the fuss is for, there is both a danger that the effort now devoted to the public understanding of science will be counterproductive and that the techniques employed will be less than effective. It is also proper to recall that, in Britain, the movement has its roots in the early Thatcher years, and in the shock with which the research community discovered that its pleas on behalf of

basic research were met by the blind incomprehension of ministers and civil servants. COPUS sprang in part from a wish to stimulate the support of voters for basic research.

Function

That is not, in itself, a respectable objective. Nor is it a realistic one. In the general understanding, it is rightly more significant that 'science' has given us Intel's Pentium chip (with all its faults) and a genetic screen for inherited cystic fibrosis than that Fermilab has found the top quark. The applications of science most directly change people's lives and naturally are given the most attention. Moreover, an academic qualification in electronic engineering is not essential for the appreciation that microprocessors are immensely useful. Most voters in most countries would probably echo their increasingly utilitarian governments in holding that the function of the scientific enterprise is to generate useful devices and techniques. Public understanding will not by itself generate a constituency for basic research.

There is no shame in that, but there is an awkward trap for those who mount public understanding campaigns: the temptation is to suppose that those asking for deeper understanding look for instruction of the kind offered to students — in this case, to students ill-provided with elementary knowledge. Too many well-meant efforts at public understanding are, as a consequence, both patronizing and unenlightening. (If in doubt on that score, try explaining why silicon rather than, say, aluminium is used in the construction of microprocessors.) Yet there is ample evidence that people with no background in science can readily acquire a good grasp of the benefits and the limitations of applications that concern them directly. The front-line troops in the battle for public understanding include such people as genetic counsellors; they manage to help people offered prenatal genetic diagnosis without going into the chemistry of DNA. The public of 'public understanding' is probably much better informed about the parts of science that matter personally than the research community allows. It needs encouragement, not instruction.

Alienation

Another stimulus of Britain's COPUS, and of similar operations elsewhere, was the fear of alienation between the research community and the rest of society. The fear is real, but the grounds for it are, or should be, spurious. The modern world is a scientific world. The gadgets that fill people's



lives (and the power sources that drive them) are often sophisticated devices at the leading edge of engineering and technology. Their economic importance, both to manufacturers and users, is immense. It would be no more feasible to dispense with them than with social institutions such as the police, a banking system or an elected legislative assembly. Yet public arguments about the role of the research enterprise are most often cast in terms suggesting that science is not part and parcel of the modern fabric, but rather a kind of elective ideology, on a par with, say, 'communism'.

One of the goals of the campaign of public understanding must be to change these terms of debate. The nature of the problem is neatly illustrated by the way in which even responsible television organizations deal with claims on behalf of 'alternative science'; they give equal time to the orthodox and the unorthodox and, at best, leave it at that. (In this spirit, the contribution of the British Broadcasting Corporation's respected *Horizon* programme to British Science Week on 20 March was a long filmed worry whether access to the Internet and its successors will be globally equitable.) That habit, which substitutes public confusion for public understanding, is bad enough, but the more serious damage is done when civil servants, consulting on legislation or other actions, also give equal weight to received opinion and its many alternatives. Only plain speaking by ministers and by the scientific community itself will remedy that state of affairs. COPUS may yet have to roll up its sleeves.

But what the experience of the past decade has shown is that the practical purpose of public understanding is to give young people an enthusiasm for science. That is not (or should not be) a surprising discovery. For nearly 20 years, Fermilab in Illinois (where the top quark was confirmed the other day) has been offering high-school students and their teachers a chance to spend Saturdays working on some physics project, explicitly to make up for the poverty of teaching in most US high-schools. The example has now been copied at other national laboratories in the United States, and even by some private corporations. And that, in reality, is what British Science Week seems to aim at.

It is an entirely proper goal. Indeed, given the continuing muddle over the pattern of education in British secondary schools, it is essential. Students wishing to study science at a university are still compelled to demonstrate, while still at school, that they know most of it already (which is a recipe for driving people into other fields). And education ministers continue to insist that this system is 'the jewel' of British education. That is why some means of rescuing able people from accountancy and other such occupations is essential. It is unlikely that British Science Week, patchily spread about the United Kingdom, will suffice to broaden young people's outlook and give them a liking for the notoriously hard grind of university science. A formal evaluation would be interesting. Meanwhile, COPUS might usefully point out that OST's sponsorship of these events is really intended to make up for the failings of the Department for Education, and that root and branch reform of the interface between schools and universities would be more effective. □

Glaxo and Wellcome

The latest pharmaceutical merger is necessarily a risk but is also an opportunity.

LARGE and successful pharmaceutical companies do not command sympathy in the ordinary sense, but their managers are no doubt acutely aware of one certainty in their working lives: the mistakes they make will be big mistakes, financially and otherwise. Sadly, at least for those with insomniac tendencies, there is another perennial truth: caution marks out the road to failure. Not to invest in the search for innovative drugs is suicidal. With the pattern of health care in many parts of the world changing to allow for direct negotiation between insurers (standing proxy for potential patients) and the providers of health care (hospitals and physicians), there are also temptations to what, in other fields, would be called vertical integration. The drug company as the comprehensive provider of health care is not entirely a fanciful concept.

That is the background to the purchase by the British drug company Glaxo of the Wellcome Foundation (not to be confused with the Wellcome Trust, which is a charity). The purchase price, a cool £9.3 billion, is one measure of the magnitude of the risk. Another, less tangible, is that the two companies have very different styles. Glaxo has grown quickly and not always smoothly over the past quarter of a century. Wellcome by contrast has been more staid, perhaps by being shielded from the rigours of full-blooded competition by the circumstance that its sole shareholder until a few years ago was the Wellcome Trust. Making the best of the two groups of people will be a challenge.

Since the possibility of a merger was first announced, there has been endless speculation about the end result, but it is probably beside the point. True, Wellcome is best known for its antiviral compounds (of which AZT is one), Glaxo for its treatment for stomach ulcers. But it is impossible for outsiders to guess meaningfully at the long-term pattern of the merged company's research. Indeed, Glaxo is fond of boasting that its research pattern is as catholic as it could be, citing its basic-research laboratory at Geneva as proof. Glaxo-Wellcome, which seems not for now to be tempted towards vertical integration, should be large enough to cover a broad portfolio of research projects.

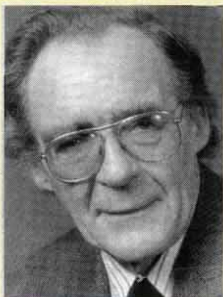
But in the modern search for new drugs, the opportunities expand faster than the companies can grow. The various human genome projects will spawn endless opportunities, for example, but only labour-intensively. Identifying a gene is one thing, but telling its function or functions is at least one and perhaps several orders of magnitude more difficult. That is why the House of Commons the other week (see *Nature* 374, 6; 1995) should not have been alarmed that the merger will reduce research employment. With luck, the trend could go the other way. The best arrangement would be a network of small research companies each working on a hunch and a handful of genes. Many of them, of course, would be academic departments. □

Maddox to step down from editor's chair at *Nature*

London. Sir John Maddox (below), editor of *Nature* for 22 of the past 29 years, will retire "on or around" his seventieth birthday in November.

Macmillan Magazines Ltd, the publisher of *Nature*, says that it has already held discussions with a number of potential successors. The company hopes that Maddox will continue to be involved in the development of the journal after his retirement.

Maddox was first appointed editor of *Nature* in 1966, having started his professional career in 1949 as a lecturer in theoretical physics at the University of Manchester. He then became science correspondent of the *Manchester Guardian* and, later, coordinator of the Nuffield Science Teaching Project.



During his first spell as editor, the circulation of the journal doubled, to 22,000. He was director of the Nuffield Foundation from 1975 to 1980, when he returned to *Nature* and has since helped to build the number of subscribers — in particular by substantially increasing its international circulation — to 56,000. He was knighted in the New Year Honours list in January this year.

"When John Maddox goes, it is bound to seem like the end of an era, but everyone hopes that his ongoing involvement [with *Nature*] will make the transition a seamless one," Nicholas Byam Shaw, the chairman of Macmillan, said on Monday.

Maddox says he intends to stay "very active", and, while maintaining links to the journal, he is also looking forward to doing "something quite different". He is currently writing a book on what science has yet to discover — and says that "it might even be fun to teach a physics course again".

Maddox says that his main regret on vacating the editor's chair is that because it has provided him with "an automatic way of keeping up to date" he will have to work harder. In addition to his academic and journalistic career, he is an elected member of the Gwenddyr and Crickadarn Community Council in Powys, South Wales. □

State university system sees 'devastating' cuts

Washington. Researchers at the State University of New York (SUNY), the largest public university system in the United States, are bracing themselves for budget cuts of up to one-third which, they fear, could undermine the university's status as a major research institution.

The threat to the US\$340 million research programmes carried out each year at SUNY has received less publicity than other threats facing the system's 400,000 students — and their parents. But faculty members at the university's major research campuses, at Buffalo and Stony Brook, expect that the anticipated cuts will hit research by raising the cost of taking on graduate students, preventing good staff from obtaining tenure and encouraging a flight of existing staff from the system.

George Pataki, the recently elected Republican governor of New York state, has put forward a budget that would drastically cut the state government's direct contribution to SUNY, reducing it from \$970 million to \$680 million next year. The governor is believed to have hit the university hard because he feels that, in contrast to other 'services' such as public schools or the sewage system, the university can raise revenues by increasing student tuition fees.

Once all revenue sources have been included, SUNY's income is about US\$3 billion a year. The proposed US\$290 million cut therefore represents a reduction of 10 per cent. Last week, Frederic Salerno, a telecommunications executive who chairs the SUNY board of trustees, said the cuts would "devastate" the university. "Another

course of action must be found," he said.

Pataki is expected to agree a budget with the Democrat-controlled state assembly and the Republican senate by 1 April. The state has previously allowed this deadline to slip for weeks or even months. But Pataki says he will shut down the state government if he does not get a budget by the agreed date, and the final figure is expected to be close to the governor's proposal.

SUNY's strongest political support has traditionally come from the Republican senators who represent upstate New York, where many the system's 60-odd campuses are located. But party loyalty has obliged them to support Pataki's budget in public. It has been left to Democrats such as Steven Englebright, a Long Island assemblyman who is also a geologist with SUNY at Stony Brook, to attack it. "This would destroy the university," says Englebright, adding that the cuts will encourage the best existing faculty to leave, and make it impossible to recruit replacements.

SUNY's \$340 million research programme makes it one of the largest research universities in the United States, spending about the same as the University of Michigan, Stanford or Cornell (although the University of California system spends more than twice as much across all its campuses). Half of SUNY's \$340 million goes on biomedical research, and most of the rest on chemistry, physics, engineering and environmental sciences.

Staff say that, like public university systems in Massachusetts and other northeastern states, SUNY suffers a widespread ►

Congress promises a bleak future for research

Washington. The current squeeze on research budgets is unlikely to be a short-lived affair, if a committee of the US House of Representatives has its way. Last week, the House Budget Committee voted to cut US\$100 billion from federal agency budgets over the next five years, in order to off-set promised tax cuts.

To demonstrate that such a reduction is feasible, the committee has published a 45-page list of what it calls "illustrative Republican spending cuts". Although these are only suggestions — the budgets of individual agencies are decided by committees in the House and Senate — they show which programmes Republican leaders believe to be expendable.

Among the recommendations are a 5 per cent reduction in research funding for the National Institutes of Health, and scaling back and delaying the National

Aeronautics and Space Administration's proposed Earth Observing System (EOS). Some of the ideas, such as eliminating the Department of Energy and the National Biological Survey are not new. But committee members seem to have changed their minds about an earlier suggestion; rather than scrapping the US Geological Survey complete, the agency is now only targeted for "significant reforms".

Appropriations committees are, in principle, free to follow these guidelines if they choose — or to ignore them. But if the \$100 billion reduction is approved by Congress within the next few months, the committees will have to find the cuts somewhere. And the budget committee plans another five-year spending reduction — this time because of the federal deficit, and up to four times as large — in May.

Tony Reichhardt

perception that students go to private universities if they can. In other parts of the United States — notably in the midwestern states — public universities are held in greater esteem and therefore they have wider political support.

Jim Kalas, associate provost for research at SUNY, says the cuts will hit research in three ways: the general atmosphere of retrenchment will encourage the best faculty to leave; \$27 million used to support graduate students will fund fewer students at higher tuition fees; and there will be no money for new equipment and buildings.

The SUNY board of trustees has drawn up a list of campus and programme closures to meet the Pataki cut. But it has not released it, on the grounds that closure threats would become self-fulfilling. Two-thirds of the \$290 million would be found, the board said, by raising tuition fees by 60 per cent.

By publishing its views on the governor's proposal last week, the board hopes to maximize its leverage in the budget process. Its fears received front page coverage in the state's largest newspaper, *Newsday*. But Pataki appears to be unmoved, and SUNY officials are privately resigned to a cut of at least US\$200 million. **Colin Macilwain**

... as NSF asked how knife could fall

Washington. The US National Science Foundation (NSF) and its governing board have been asked by a House of Representatives subcommittee how it might absorb a 20 per cent cut in its US\$3.3 billion budget for the fiscal year 1996, starting on 1 October.

The request was made last week by Jerry Lewis (Republican, California), chairman of the House appropriations subcommittee for Veterans' Affairs, Housing and Urban Development and independent agencies (VA-HUD), to Neal Lane, director of NSF, and Frank Rhodes, chairman of the National Science Board.

Lewis is not yet proposing a 20 per cent cut at the NSF, he said later, but he is asking the foundation to say what it would do to achieve such a cut. Similar requests were being made to other agencies covered by the subcommittee, which is responsible for \$70 billion — more than a quarter of all non-military money controlled by Congress. Earlier, he had pointed out that NSF's budget had grown by 10 per cent since last year, but he also expressed concern that "very often, science is an easy target for cuts".

In reply, Rhodes told Lewis that the National Science Board, which meets this week, was prepared to help the House set priorities for the NSF. But he said later that he had interpreted Lewis's request as concerning options for 20 per cent savings spread over several years.

But Lewis and his staff say they want ideas on how such a cut could be made in the 1996 budget — and that they need pro-

NIH to review priorities in readiness for future cuts . . .

Washington. Faced by inevitable cuts in spending over the next five years — regardless of which political party is in power — the US National Institutes of Health (NIH) are to carry out a wide-ranging review of biomedical research in the United States.

"We will talk to all the universities, to find out how they want biomedical research to look by the year 2000," Harold Varmus, director of the NIH, said in testimony before a congressional committee last week.

Varmus was responding to remarks by John Porter (Republican, Illinois), chairman of the House of Representatives appropriations subcommittee on labour, health and human services, who had just explained that increased spending on health-care and welfare is squeezing out even the highest priority discretionary programmes.

The NIH budget, which stands at \$11.2 billion this year, with President Bill Clinton having requested \$11.8 billion for next year, is part of this discretionary spending. Last

week's hearing was the first of many at which the NIH will defend the new request.

There is support from both Republicans and Democrats for the NIH's budget, and from Porter in particular, and the institutes are unlikely to be in a worse position under the Republicans than they were before. Porter's subcommittee organized a hearing two weeks ago at which a number of Nobel prize-winners explained the advances that they had made with government support, and the significance of their work to the well-being of the United States today.

The hearing was held partly to put on the record arguments that Republicans wishing to defend the NIH budget will be able to use. But it also gave Porter an opportunity to explain the exigencies of the current budget crisis. He repeated these last week, prompting acknowledgement of the situation from Varmus and his promise that the NIH would examine how biomedical research should accommodate cuts in funding.

Last week, the House Budget Committee entered the fray with its proposal to cut \$100 billion from discretionary funds at 1995 levels over the next five years (see previous page). The NIH would decide the priorities for these cuts, the committee said.

But such priorities would need to be defended. And as Varmus discovered last week, when it comes to specific programmes, agreement between the parties can rapidly break down. Jay Dickey (Republican, Arkansas) asked Varmus how he justified "research into lung cancer and AIDS, diseases brought on by people's own behaviour". Dickey said he needed such justification for the voters in his home constituency.

Varmus argued that behaviour can play a part in most diseases. But Dickey replied that this was not the case in rheumatoid arthritis, and then asked what the NIH was doing about heart disease, an area of research he clearly supports.

Led by Henry Bonilla (Republican, Texas), the subcommittee then became embroiled in a discussion about how many people are killed by AIDS in comparison to other diseases. Bonilla eventually asked Varmus why AIDS is so contentious both among the public and among researchers; it took Varmus a moment of reflection to understand the question, which he then carefully circumnavigated.

To judge by what subcommittee members said, the hearing suggested AIDS research may run into opposition. Yet most of the committee had voted to restore funding for AIDS research cut from the 1995 budget earlier this year. Last week's hearing was the beginning of a long road; as yet, it is hard to see where it is going. **Helen Gavaghan**

Colin Macilwain

Japan placates fishermen after launch delay

Tokyo. Japan's ambitions to become a major player in commercial space activities received a boost last Saturday (18 March) with the successful launch of two large payloads by its new H-II rocket. The launch, which had been delayed several weeks because of technical problems, will help to restore confidence in the H-II, which was dented by the failure of its second mission last August (*Nature*, 371, 92: 1995).

The rocket successfully placed a geostationary weather satellite in transfer orbit. It also released a reusable orbital platform, the Space Flyer Unit (SFU), which has been designed to carry out various scientific experiments before being recaptured by the US space shuttle in November.

The SFU, which cost ¥41.8 billion (nearly US\$0.5 billion) to develop, is a huge joint mission involving three different government organizations, and thus represents a rare example of collaboration in research between different arms of the government.

The agencies involved are the Institute of Space and Astronautical Science (ISAS) of the Ministry of Education, Science and Culture, the National Space Development Agency (NASDA), which is affiliated with the Science and Technology Agency, and the New Energy and Industrial Technology Development Organization (NEDO) which operates under the powerful Ministry of

International Trade and Industry.

The platform was due to be boosted on 21 March from a height of about 330 km to a higher operational orbit of 500 km. It carries an exposed section on which NASDA is planning to test equipment and materials intended for the Japanese Experiment Module of the US space station.

ISAS, which is responsible for tracking the SFU, has several experiments on board. These include an infrared telescope, a high voltage solar array designed to verify that changing the connections between solar cells can alter its output, and an experimental electric propulsion system called the Magneto-Plasma-Dynamic thruster.

Other experiments include a facility for crystal growth experiments under conditions of microgravity, and a container with newts that will be used to test the effects of zero gravity on egg laying and hatching. NEDO has three electric furnaces aboard the platform to grow semiconductor crystals and thin films under microgravity conditions.

The launch was originally planned for the beginning of February, but experienced a

series of delays because of problems with the SFU and other components. As a result, NASDA was put in the unusual position of having to negotiate with local fishermen to allow the launch to proceed.

This resulted from an earlier agreement reached with about 10,000 fishermen in five prefectures adjacent to NASDA's launch centre in Tanegashima island off Japan's southern island of Kyushu. Under the terms of the agreement, NASDA pays about ¥600 million (\$6 million) a year to fishing cooperatives for permission to launch during two two-month periods, in summer and in winter.

As this week's launch was outside the January–February window, NASDA had to strike a special one-off deal with the fishermen, who claim the launches disrupt their activities. NASDA will not say how much was paid, as it varied from cooperative to cooperative. The restriction on two launch windows a year remains, and the agency will have to reach a long-term agreement with the fishermen if it is to become a serious competitor in the world market for commercial satellite launches. **David Swinbanks**

IMAGE
UNAVAILABLE
FOR
COPYRIGHT

Japan's Space Flyer Unit: a successful example of inter-ministerial cooperation.

Grant challenges role of peer review in an age of e-mail

London. The US National Science Foundation (NSF) has awarded a grant of more than US\$1 million to Paul Ginsparg, a high-energy physicist at the Los Alamos National Laboratories in New Mexico, to expand his electronic archive of research articles and preprints.

Ginsparg set up the archive, initially specializing in high-energy physics, in 1991. It now includes 25 other disciplines and is consulted more than 45,000 times a day by scientists worldwide, wishing to find or to contribute new items. But it has also raised controversy by challenging the conventional practice of publishing papers only when they have been peer-reviewed.

Researchers are able to send in articles or preprints in relevant subject areas. Once the papers are in electronic format, Ginsparg provides cross-references. He says that his archive differs from computer bulletin boards and news groups in that it is a "formal mode of communication in which each entry is archived and indexed for later retrieval".

Richard Isaacson, programme director for gravitational physics at the NSF, describes Ginsparg's database as "an unusual and unique service" that has "already become the most interesting stop

on the infobahn for a large part of the theoretical physics community".

The purpose of the grant, which has come from both the NSF's physics budget and an Office of Multi-Disciplinary Activities created last November to support innovative projects, is to allow Ginsparg — who has so far been developing the database on a part-time basis — to take on full-time professional software support.

The main novelty of Ginsparg's system over the conventional form of informal electronic communication between scientists is that it allows public commentaries to be added to papers, an approach that could take the place of traditional refereeing.

Some see this as a direct threat to the peer-review process used by traditional printed journals. "Publishers might not like what is happening," Ginsparg acknowledges. "But the effort to change the manner in which research is communicated is being driven by the researchers."

Most academic publishers are already anticipating a gradual shift from paper to electronic archives of publications. But there is still much heated debate about the

implications of the traditional process of refereeing and editing papers before publication being replaced by Ginsparg's vision of "interactive journals".

Ginsparg himself says that he hopes his archive will support what he describes as a "virtual corridor", where a scientist will be able to work down the hallway and talk to fellow researchers.

At the same time, publishers themselves are considering how to build an electronic infrastructure to suit researchers' needs. Last month, for example, the publishing company Elsevier Science launched a pilot scheme to investigate how electronic subscriptions can be used.

Roland Dietz, senior vice president of Elsevier Science, New York, says that research institutions in both the academic and industrial sector are at a wide variety of stages of development. "Some already have internal systems in place, others are much less advanced."

But Dietz says that all recognize that the potential is enormous. He says that electronic libraries will be able to store information locally, make it available to researchers at their desktops, set up security systems and allow users to log on from home or a conference. **Fiona Gammie**

France joins Italy in mouse gene store bid

Rome. Mario Capecchi, an Italian-born geneticist who has played a major role in the development of knock-out gene technology — a technique by which animals are genetically engineered to contain a non-functioning copy of a specified gene — is to join the staff of a new mouse genetics centre being set up by the Italian government at Monterotondo near Rome. Capecchi is currently working in the United States.

The appointment will strengthen the centre's bid to host a planned European mouse mutant repository, a collection of mutant strains of mice, kept either alive or as frozen embryos, which will be made available on request to research scientists (see *Nature* 372, 305; 1995).

Given the increasing importance of such mice in research, the European Commission (EC) has set aside special funds in its fourth Framework programme for mutant repositories. Research organizations in Italy, France, Germany and the United Kingdom are backing an application to house the mouse gene repository at Monterotondo.

Despite a general freeze on public appointments, Italy agreed last week to release money for Capecchi's appointment from a scheme set up by the former research minister, Antonio Ruberti, to bring leading scientists working abroad, back to Italy.

Capecchi's role has not yet been precisely defined. But it will involve an overseeing role that will include coordinating the interaction between the international and national activities of the centre.

Capecchi, who is 57, will retain his post at the Howard Hughes Medical Center in Utah. He says that at first he will simply be giving advice on setting up the centre. Later, however, "there is a good possibility of going beyond this". He says that he will probably spend more time in Italy once the centre is operating, and looks forward to doing what he can "to help continue Italy's tradition of producing good scientists". Although Italy remains strong in physics, its expertise in biology has declined over the past two decades, something that it is keen to correct.

Attracting Capecchi to the centre adds weight to its application to host the EC's main mouse mutant repository, as a strong research environment is considered essential. The Consiglio Nazionale della Ricerche (CNR) has already agreed to move its own Rome-based molecular biology research groups of about 100 staff into the Monterotondo laboratories, which were donated by the Italian ENI oil company. In addition, four regional groups of the European Molecular Biology Laboratory (EMBL) will be established there by the end of the year.

According to Enrico Garaci, the head of the CNR, the combination of these activities means that Italy will have a strong research environment at Monterotondo to back its

application to Brussels. The CNR has already agreed to finance the expansion of the genetic repository facilities if this becomes necessary after a few years.

Consensus between the CNR, Germany's Max Planck Society and France's Centre National de la Recherche Scientifique (CNRS), which last year set up a joint search committee to identify a suitable site for the proposed mouse genetics repository, has not been achieved without a struggle.

The committee wanted a main laboratory with a series of "networking nodes" in other countries to provide additional input or services to back up the main facility. The nature and relative importance of the nodes was initially a point of heated argument between the organizations.

France, which recently complained that Italy would win the bid for the host site because of no competition (see *Nature* 373, 276; 1995), was particularly keen for its own CNRS animal breeding laboratory in Orléans to play a significant role.

The French view has now prevailed. In the application, submitted before the deadline this week, the CNRS laboratory would be the main supporting laboratory — or

'node' — which would duplicate the frozen embryos stored at Monterotondo. "For security reasons, [the CNRS] thought it would be good to have at least two places where frozen embryos are kept," says François Galibert of CNRS.

The new plan is also of direct benefit to France. Two years ago, CNRS decided to expand its animal facilities in Orléans to act as a national mouse mutant repository. The laboratories look after around 150 live strains of transgenic and knock-out rodents, and around 250 strains as frozen embryos. But its small research facility will be joined later this year by a larger research group transferred from Paris, and more groups will be encouraged to move there later, according to Galibert.

If the EC application covering a main laboratory at Monterotondo and a main node at Orléans — costing just under ECU6 million (US \$4.6 million) over a four-year period — is successful, additional national nodes will be added later. A second phase application to the EU, that is likely to include a bid for another 'node' from Britain's Medical Research Council, will be submitted by the end of the year.

Alison Abbott

UK puts science in the limelight

London. **Sunday morning at London's Euston railway station. All is quiet except for the arrival (left) of a 15-foot replica of the DNA double helix, not part of this year's charity fund-raising extravaganza "Comic Relief", but of another deserving cause: promoting British science.**

The model, designed by Alan Ward, forms the centrepiece of an exhibition about modern genetics organized by the Wellcome Trust and the Medical Research Council, one of over 3,000 events celebrating the second UK National Science Week.

Money has been provided for individual events by a mixture of private sponsors including British Gas and IBM, who sponsored the opening ceremonies, grants worth £95,000 (US\$ 150,500) from the Committee on the Public Understanding of Science.

As a reward for the £50,000 grant from the Office of Science and Technology, John Horam, the junior science minister, received a guided tour of the microbiology of brewing courtesy of the 'yeast beast' — "a giant flocculating head of beer with a demented look on his face" — at the Bass [Brewers']

**IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS**

Museum in Derbyshire.

This year has seen an attempt to give Science Week a more serious flavour with press briefings being held by all the research councils, in addition to public events. The Health and Safety Executive used the week as a platform to launch its former Research and Laboratory Services Division as a new government agency, the Health and Safety Laboratory.

David Hunt, the science minister, gave a keynote address on 17 March. He told his audience the date was a particularly appropriate one on which to open the Science Week as it marked the anniversary of a major (and patented) British invention: the elastic band.

F.G.

Universities in clash over Internet 'libel' allegations

London and Montreal. A British academic, stung by a series of personal attacks posted on the Internet, has threatened to sue six North American universities for their part in 'promoting libel'. In what could be the first case of its kind, Laurence Godfrey, a 41-year-old London-based physics lecturer, has written to five universities in Canada and one in the United States asking them to censor allegedly libellous messages about him mailed from their campuses to the user group Usenet.

Godfrey, who worked as a high-energy physicist for the National Research Council (NRC) of Canada between 1984 and 1992, says universities must bear responsibility for electronic mail passing through their computer terminals. "Universities think I am threatening freedom of speech. But what about the damage done to my reputation?"

He claims that the material transmitted "has lowered my esteem in the eyes of students, colleagues and future colleagues". Godfrey adds "if a radio station can be sued for broadcasting libellous remarks on a phone-in programme, the same must apply to institutions linked to the Internet."

The statements, which first appeared in 1993 and were reissued from several sites, refer to his departure from the NRC. Internet users in Canada claim that the British physicist baited Canadians by suggesting that theirs is a country of third-rate people doing third-rate work, and thus drew inflammatory responses.

The universities have indicated that they intend to defend any proposed action, despite having paid out compensation in an out-of-court settlement. The Canadian Universities' Reciprocal Insurance Exchange, which is handling the claim, is now demanding a refund. It complains that Godfrey breached the terms of settlement by, ironically, announcing it on the Internet.

Universities tend to agree with Internet user groups that interference with Internet messages is a violation of free expression, and they do not consider themselves responsible for information fed into the network through their terminals.

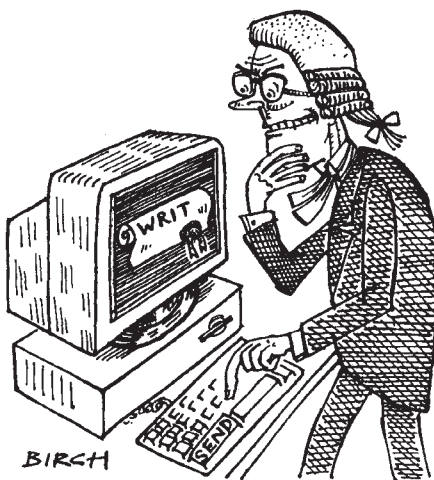
One university, which cannot be identified for legal reasons, says that any talk of censorship of the Internet conflicts with the ethos of allowing unrestricted freedom of speech. The question of whether a particular message is libellous, the university adds, can be decided only by the courts.

Legal experts are reluctant to speculate on the outcome if Godfrey decides to sue. Nick Braithwaite, a media lawyer with Clifford Chance, a firm of London solicitors, says legislation governing the issue is "hazy and unclear".

In principle, universities are protected by the so-called Innocent Dissemination Act, a piece of legislation designed to safeguard organizations from being sued for defamatory material illegally or unwittingly issued in their name.

But this act could work in favour of both sides. The universities will benefit if the courts decide that the material relating to Godfrey was not defamatory. But Godfrey is counting on the fact that the universities are no longer 'innocent' parties, in the legal sense. "It is not a case of innocent dissemination any more," he says. "The universities know that false information is being put out through their channels, yet they still refuse to take action."

The government, according to Nick Braithwaite, is in the process of reviewing



the Innocent Dissemination Act. "It is up to universities to lobby hard and clear up ambiguities," he says.

Meanwhile, they should ensure that they have insurance cover for third-party liability and should issue guidelines warning of the legal pitfalls of bulletin-board use. "People can still have private conversations on the Internet," he said. "But they need to be careful about what they say when using bulletin boards like Usenet and Compuserve."

Godfrey is already pursuing another lawsuit that involves the Internet. This is a case against Phillip Hallam-Baker, a researcher at the European Laboratory for Particle Physics (CERN) in Geneva, which is due to be heard in the high court in London in July. Godfrey, who previously worked at DESY, the German Electron Synchrotron Laboratory, in Hamburg, claims that Hallam-Baker made damaging remarks on Usenet about his professional work. Hallam-Baker says he is ready to defend any action.

David Spurgeon & Ehsan Masood

Wellcome pledges \$80 million for population studies

London. Britain's Wellcome Trust, which last week became the world's richest research foundation, has announced that it is to spend £50 million (US\$80 million) over the next five years supporting studies of ways to curb unwanted population growth.

Much of this money will be spent outside the United Kingdom, reflecting the trust's efforts to increase the international scope of its activities. Two other enhanced priorities will be its support of studies of research-funding decisions, and of moves to stimulate the development of new technologies based on the work of Wellcome-supported scientists (see *Nature* 372, 6; 1995).

The trust's announcement of its new priorities coincided with the formal acceptance last week by the Wellcome Foundation — the pharmaceutical company whose shares were vested in the trust on the death of its founder, Henry Wellcome, in 1936 — of a takeover bid by its larger rival, Glaxo.

The bid was successful largely because of the trust's early agreement to sell to Glaxo the 40 per cent of Wellcome shares remaining after two earlier divestitures. It has increased the trust's total 'investment base' to £6.7 billion, and thus its anticipated annual income by an estimated £50 million to a sum approaching £300 million.

But according to Bridget Ogilvie, the director of the Wellcome Trust, its new priorities were in the pipeline before there was any knowledge of the Glaxo bid. "For example, we have been progressing towards the extra funding for population studies for some time," she says, pointing out that one spur was last year's UN Conference on Population in Cairo.

The new directions are also intended to build on developments since the last major sale of Wellcome shares, in 1992. Since then, the trust has concentrated primarily on building up support for research groups in British universities through fellowship schemes, equipment grants and the funding of new buildings.

"We have not yet focused on developing our international portfolios, but this is now our intention," says Ogilvie. Similarly, on the question of technology transfer, she says that the trust is on "a very steep learning curve; we are very keen to get moving".

Meanwhile, the House of Commons Select Committee on Science and Technology published a report on 20 March, based on a meeting with senior officials from Glaxo, the Wellcome Foundation and the Wellcome Trust last month (see *Nature* 372, 6; 1995), saying that "only time will tell" whether the merged company will be as strong as its chief executive, Sir Richard Sykes, is predicting. □

Europe extends biotech grants 'experiment'

Paris. On the basis of the apparent success of a pilot project, the European Commission is to delegate to the scientific community the management of four biotechnology projects funded by the European Union (EU)'s Fourth Framework research programme. But the body responsible for the pilot project says it no longer proposes to carry out the initial selection of research proposals.

The commission's move extends an experiment started two years ago within the third Framework programme, under which a consortium known as A Molecular Initiative in Community Agriculture — or AMICA — has managed an ECU24 million (US\$18.5 million) Priority Technological Programme (PTP) on Plant Molecular Genetics for an Environmentally Compatible Agriculture (see *Nature* 366, 291; 1993).

The consortium is led by the John Innes Institute in Norwich (United Kingdom) and the Max Planck Institute in Cologne (Germany). According to one commission official, it is too soon to say whether delegation of the programme's management has improved its scientific output. But from an administrative point of view the test case has "worked well and quickly" in getting grants to scientists and recruiting staff, he says.

Many researchers agree. "AMICA's management has been very efficient," says Jean Dénarié, a researcher at the CNRS/INRA Molecular Biology of Interactions between Plants and Microorganisms laboratory in Toulouse.

Another advantage, say researchers, is that whereas EU programmes have previously consisted of many small consortia — usually made up of five laboratories — AMICA coordinates all 117 laboratories taking part in the PTP within four 'networks' grouped around 18 research themes.

On the basis of such positive responses, the commission says it intends to extend the AMICA model of delegated management to the successor to the current PTP, the plant and animal biotechnology programme. It will also be applied to three other projects within the biotechnology programme: the cell factory, genome analysis and cell communication in the neurosciences.

A bid from AMICA to run this next plant biotechnology programme would be "most welcome", says one commission official, adding that other bids would also be considered for both this and the other programmes following an open call for proposals which the commission will make in the autumn.

Jeff Schell, director of the Max Planck Institute, says that AMICA will apply to run the new programme, but it will not propose that the consortium should also carry out the initial peer review and selection of pro-

jects. An attempt by AMICA to do so in 1993 provoked controversy, while half of the projects it selected were eliminated by the EU's own peer-review process.

Schell admits that there was much opposition to AMICA's attempt to take responsibility for selecting projects. But he says its decision not to apply for this role in the next Framework programme was taken only because it would be unable to "do the job well" within the available time. "We don't want to lose our credibility, and play into the hands of critics," he says.

Nevertheless, Dénarié and other scientists are concerned about the "principle" of allowing consortia of scientists to select projects. "It would be unimaginable for the National Science Foundation (NSF) to hand over selection of grants to a small committee made up of representatives of just several US universities."

Declan Butler

Germany defuses conflict over labs

Munich. A sharp conflict about whether industry should have more control over the work of Germany's state-funded national research centres was defused last week at a meeting between government officials, directors and union representatives from two major national research centres, and industry representatives.

In 1993, the ministry of research and technology — now known as the Bundesministerium für Bildung und Forschung (BMBF) — commissioned a study of industrial relevance of the research at the national centres.

The study was conducted by a group led by Hartmut Weule, head of the research division of Daimler-Benz, and focused on

the two largest of Germany's 16 national research centres, the Forschungszentrum Jülich (KfA) and the Kernforschungszentrum Karlsruhe (KfK), both of which have applied and basic research programmes.

It found about 30 per cent of work in the two centres was relevant to industry — and recommended that this be increased to 75 per cent over a five-year period (see *Nature* 372, 4; 1994). But the Weule commission is now backtracking on the 75 per cent figure; Wolfgang Scheunemann, a spokesman for the commission, says that it was not intended literally.

It has now been agreed that communication between the two sides should be increased, in particular at the planning stages of research projects. Joachim Treusch, KfA director, says the number of industrial scientists on advisory committees, as well as their voting rights, has already been increased at his institute, and will be further increased.

But behind the publicly expressed harmony, it remains unclear whether the industrialists and the national research centres really see eye-to-eye on what they each mean by "strengthening" applied research. The two national research centres agree that they need to strengthen certain applied research programmes, but not necessarily that they should increase the total number of applied research programmes, says Treusch.

Additional measures to increase collaboration between industry and the national research centres, such as exchange of scientists between the two, are being discussed. A further meeting is scheduled in May.

Toni Feder

India keeps the squeeze on science budget

New Delhi. For the third year in a row, money for research and development will be scarce in India's government laboratories and universities next year, as the budget for 1995–96, published last week, provides only a small increase in funds for science.

The 37.7 billion rupees (US\$1.2 billion) allocation — omitting 9.4 billion rupees spent on defence research — is 3 billion rupees more than last year. But the increase is likely to be less than the rate of inflation, now running at 11.5 per cent.

Pleas by prominent scientists for more research and development funding appear to have gone unheeded by the government, whose term of office ends next year. A large proportion of the budget has been allocated to populist programmes such as alleviation of poverty, rural development and benefits to farmers — all likely to boost the ruling party in the general elections.

Apart from a token grant for the proposed superconducting LINAC booster for research in heavy-ion nuclear physics, the only new project for which money has been approved this year is a DNA fingerprinting centre, to be set up at Hyderabad. The centre will provide fingerprinting services for both criminal investigations and settling paternity disputes, using an indigenously developed DNA probe.

Unlike the previous two years, the budget squeeze will have no adverse effect on the Council of Scientific and Industrial Research (CSIR), India's largest research agency with about 40 laboratories, which earns substantial funds from industrial contracts.

But other agencies will not be so lucky, and will receive only enough funds to maintain existing projects. "There will be a complete slackening of research in universities," says one scientist.

K.S.Jayaraman

Italy widens peer review of health research

Rome. Italy's largest health research institute, the Istituto Superiore di Sanità (the Higher Institute of Health) has won the right to distribute funds for some of its research programmes, previously restricted to staff scientists, to research groups elsewhere in Italy.

Leading researchers at the institute will take responsibility for running these programmes and in particular for establishing peer review systems that will ensure grant applications are approved solely on the basis of scientific merit, rather than to meet an administrative or political agenda.

The institute was founded in 1934 by the Rockefeller Foundation, and now controls a total annual budget of IL300 billion (US\$ 173million). It has a staff of 1,400, a third of whom are scientists, and carries out regulatory and advisory activities for the government, including setting standards and monitoring the quality of blood products.

In addition, the institute receives funds to support research related to public health, following a programme agreed between the Ministry of Health and the institute's scientific committee. (Basic biomedical research in Italy is supported by the Consiglio Nazionale della Ricerche, the CNR).

Until recently, research programmes were conducted exclusively within the institute. But in 1987 a new law was passed allowing a committee at the institute, headed by Giovanni Rossi, a virologist, to run the Italian AIDS research programme. This receives around IL28 billion a year, nearly half of which is destined for basic research into the pathogenesis of HIV infection. The programme became widely admired for its efficient administration and in particular for introducing a formal system of evaluating grant applications, based on peer review by both Italian and foreign scientists.

Italy's lax standards in evaluating grants supported by public research funds, and the lack of transparency of many grant-giving bodies such as the CNR, which does not publicly state its criteria for awarding grant money, have long been a source of frustration, particularly among leading scientists (see *Nature* 367, 398; 1994).

"Universities, for example, have adopted the rain principle," says Fernando Aiuti, professor of immunology at the University of Rome 'La Sapienza', and a leading AIDS researcher. "Everyone gets a drop and they are therefore kept happy." But, he points out, no-one gets enough money from a single grant application to do significant work.

In contrast, he says, those responsible for the AIDS research programme have not been afraid to provide large sums to only a

few of the best research groups, he says. Each application is reviewed by six experts inside Italy and two from abroad. "This type of system, with referees, is the only one that guarantees honesty and high standards," comments Aiuti.

The director of the Istituto Superiore di Sanità, Giuseppe Vicari, has been keen to extend this approach to other research programmes financed by the health ministry. It



Aiuti (left): says peer review "guarantees high standards"; (above) Istituto Superiore di Sanità

took some time to convince the Court of Accounts that the 1987 law could be applied to other research areas, as the law referred only to AIDS research. But after a meeting between the court and Vicari earlier this month, the principle was accepted. "This means nine other areas, including tuberculosis, multiple sclerosis, hepatitis and gene therapy, may now be handled similarly," says Vicari.

The Istituto Superiore di Sanità has other reasons to feel optimistic about its future as a research institute because, after many years of political uncertainty, it is finally getting itself on a firmer footing.

The institute lost much of its original prestige during the 1960s, when many areas of Italian life became politically polarized. The founding director, Domenico Marotta, and his left-of-centre scientific and administrative boards, fell out of favour with the ruling Christian Democrats. Under pressure to retire, Marotta was eventually charged with misusing funds and jailed, while many scientists left the institute.

UK virology institute 'to change direction'

London. David Bishop, director of the Natural Environment Research Council's Institute of Virology and Environmental Microbiology outside Oxford, has been removed from his post. He was told last week that the research council is seeking a new director for the institute.

The institute has been at the centre of a fierce public controversy about experiments involving field trials of a potential pesticide based on a plant virus genetically engineered to express a scorpion toxin.

But NERC officials deny any link with

In 1973 a new law, intended to introduce more democracy into the administrative structures, was passed setting up an advisory scientific council and an executive administrative council, both with elected representatives from all staff levels.

But the new arrangements proved unwieldy and in 1992, as part of government efforts to make Italy's public sector more efficient, a further law was passed, allowing

changes in the national health service. As a result, a decree took effect last month reforming the tasks and structure of the institute's two councils.

The scientific council, which is required to express its views on all scientific activities at the institute, will now have only three elected members, representing research staff. All the institute's laboratory directors (around 15) will automatically be members. Nominated representatives from relevant ministries and regional governments complete the council.

One task of the new council will be to monitor potential conflict of interests over research funded from outside sources. For the first time, researchers at the institute will be able to accept private funding, including that from the pharmaceutical industry. But some staff members fear this could conflict with the main concern of the institute, namely the improvement of public health.

The issue is particularly sensitive because the previous director, Francesco-Antonio Manzola, formerly professor of anatomy at the University of Bologna, was arrested in 1993 on corruption charges relating to the fixing of drug prices. Although the institute was not associated with the scandal, its scientists are acutely conscious of the embarrassment they were caused.

The administrative council, which previously had to approve every minor expenditure, will now have less day-to-day control, and will be required only to approve a general budget annually, and to check expenditure at the end of the year. **Alison Abbott**

the controversy, which first erupted when local residents claimed they had not been properly informed about the experiments. In a prepared statement, the council said that the move followed a review carried out in line with government policy.

"It has been decided the mission of the institute will be refocused to undertake basic and strategic research to generate an understanding of the biodiversity and functional roles of microbial populations in the environment," said the statement from the council. □

Reducing greenhouse gases

SIR — Victor and Salt¹ argue that the goals of the United Nations Framework Convention on Climate Change can be advanced most effectively by focusing on improved reporting and review procedures rather than on adopting legally binding targets and timetables for reducing greenhouse gas emissions. This assertion appears to be based on the very disappointing nature of most national plans developed in response to the 'soft' target currently in the convention, which requires industrialized countries to adopt policies and measures "aimed" at returning greenhouse gas emissions to 1990 levels by 2000. The limited success of the soft-target approach, however, indicates that legally binding emission-reduction targets and timetables are needed, not that targets in general are ineffective. The remarkable success of the Montreal Protocol on Substances that Deplete the Ozone Layer clearly demonstrates that internationally agreed emission reduction requirements can work.

It is certainly true that national efforts to reduce greenhouse gas emissions have so far been extremely deficient, as documented in a recent review conducted by nongovernmental organizations in 20 OECD and 10 Central and Eastern European countries^{2,3}. This review reveals that many countries have failed to adopt politically difficult measures, at least in part because they expect that other countries will also fail to meet their commitments. Sole reliance on the report and review procedure recommended by Victor and Salt thus has the potential to lead to a race to the bottom, rather than the hoped-for positive reinforcement.

Adoption of political targets in the absence of clearly defined plans for achieving them does indeed create a serious risk of non-compliance. The solution proposed by most OECD governments at the preparatory meeting for the Conference of Parties on 9 February is to negotiate a protocol that includes both targets and agreed measures (see, for example, Statement of France on behalf of the European Union Intergovernmental Negotiating Committee for a Framework Convention on Climate Change (11th Session), New York, 8 February 1995). This approach gives each party assurances that it will not put itself at a competitive disadvantage by acting unilaterally. Reaching final agreement on such a protocol will undoubtedly be difficult, but there is no alternative if the problem of global climate change is to be seriously addressed.

Daniel A. Lashof

*Natural Resources Defense Council,
1350 New York Avenue, NW,
Washington, DC 20005, USA*

SIR — The countries negotiating targets during forthcoming meetings of the Conference of the Parties for the Framework Climate Convention should consider a broader range of ideas on commitments than have circulated in the three years since the convention's signing in Rio de Janeiro in 1992. Contrary to the argument of Victor and Salt¹, targets of some kind are necessary to drive change, as indicated by decades of experience with environmental problems both before and after targets and standard-setting in domestic, bilateral and international environmental agreements⁴. However, as Victor and Salt mention in passing, countries are not likely to meet any numerical emissions target precisely.

A review of historical CO₂ data reveals that annual emissions from energy consumption fluctuated wildly in many Annex I (industrialized) countries during the 1980s. Of 23 industrialized countries considered, about one third experienced an emissions increase of at least 10 per cent from one year to the next during the period 1980–90; emissions in all these countries differed by more than 5 per cent from a longer-term national trend in at least one year during this interval⁵. Given that CO₂ emissions are the product of many confounding forces, emissions will vary widely until greenhouse gases become as marginal (or as sectorally concentrated) as other things that are given targets in international agreements such as halocarbons, Mark 12A warheads, or even sulphur dioxide. It is the nature of democracy that national economies are inherently unpredictable, but even alternative systems cannot plan emissions exactly, as the rolling CO₂ emissions time series of formerly centrally planned countries indicate. Instead of setting only one national emissions target for one year, for example, CO₂ stabilization in 2000 at 1990 levels, parties could adopt an alternative target averaged over an interval.

A stabilization target could be accompanied by a population-adjusted average carbon intensity improvement target. As they are easier to achieve, goals for annual average change in CO₂/(GNP/capita) should be more stringent than national greenhouse gas emissions targets. For example, if countries fail to meet an emissions stabilization target in 2000, but can show that they improved their per capita carbon intensity by an average of 2 per cent each year during 1990–2000, they may still be judged to be in compliance. Other countries may aim for and reach an emissions target, but given the uncertainty in projecting economic swings, all but the least carbon-intensive countries may well decide to aim for the intensity target as backup. Historical data indicate that

carbon/output relationships show greater linearity than national carbon trends^{5,6}. An intensity target would help to focus planning on innovations in efficiency and conservation rather than relying on the 'recession' programme implicit in many nation's greenhouse gas projections. Targets based on average annual change might also lessen the temptation to inflate the baseline emissions estimates, a tactic evident in several of the 1990 fossil fuel-related emissions inventories submitted last year.

Finally, average annual targets — applied to either CO₂ intensity or CO₂ emissions changes — will help to steer climate programmes and their review towards the longer term rather than relying on the uncertain outcome of one arbitrary target year.

S. E. Subak

*CSERGE,
University of East Anglia,
Norwich NR4 7TJ, UK*

1. Victor, D. G. & Salt, J. E. *Nature* **373**, 280–282 (1995).
2. US Climate Action Network and Climate Network Europe. *Independent NGO Evaluations of National Plans for Climate Change Mitigation: OECD Countries* (Climate Network Europe, Bruxelles, 1995).
3. Climate Action Network — Central and Eastern Europe and Climate Network Europe. *Independent NGO Evaluations of National Plans for Climate Change Mitigation: Central and Eastern Europe* (Climate Network Europe, Bruxelles, 1995).
4. *OECD Environmental Performance Reviews* (Organization for Economic Co-operation and Development, Paris, 1994).
5. Marland, G., Andres, R. J. & Boden, T. A. *Global, Regional, and National CO₂ Emissions* (Oak Ridge National Laboratory, Oak Ridge, 1993).
6. *International Financial Statistics Yearbook* (International Monetary Fund, Washington, 1994).

When life begins

SIR — There still seems to be some misunderstanding about when the new life begins^{1–3}. This clearly must be when the egg is activated to carry on with its development to completion. Without activation it stays as it is, until it eventually dies. And although an egg is normally activated when or soon after it is fertilized, that is not always so. For example, frogs' eggs can be activated to develop parthenogenetically, with no sperm involved at all.

But whether or not this is of any ethical or theological significance. Godfrey is wrong in saying⁴ that "Activation, though rapid, is not instantaneous". It is true that the processes ending with activation may take seconds, or possibly minutes, from start to finish, but the actual activation itself is indeed instantaneous in the sense that the egg either is activated or it is not, and it cannot be half activated.

C. B. Goodhart

*Gonville & Caius College,
Cambridge, UK*

1. Jones, D. A. & Haynie, D. T. *Nature* **373**, 652 (1995).
2. Herranz, G. *Nature* **373**, 652 (1995).
3. Jarmulowicz, M. *Nature* **373**, 652 (1995).
4. Godfrey, J. *Nature* **373**, 100 (1995).

Questionnaires and copyright

SIR — I agree that there should not have been civil litigation in the case of Dr T.H. Lam (*Nature* 373, 458 & 465; 1995).

In 1986, when the dispute arose between Dr Linda Koo and Professor J.H.C. Ho on the one hand and Dr Lam on the other, no formal mechanism existed within the university for the redress of staff grievances. The university, under my predecessor, nevertheless made every possible attempt at conciliation: an *ad hoc* group, chaired by a pro-vice-chancellor, was set up to enquire into the matter. But Koo and Ho chose to initiate civil court proceedings.

When I arrived at the university as vice-chancellor, I offered to try to reconcile the parties. Because they had initiated court proceedings, I sought advice from the university's solicitors, who advised me that the internal enquiry should be stopped, on the grounds that neither party would want to disclose information that might adversely affect his or her own case. Since then, we have instituted formal procedures, although they are as yet untested.

Wang Gungwu

*University of Hong Kong,
Pokfulam Road, Hong Kong*

SIR — I represented Dr Lam both before the Court of Appeal in the case brought by Dr Koo and before the Hong Kong University Committee on Personal Matters, a disciplinary committee of the university, and the question of whether Lam had "copied" any part of Koo's questionnaire was extensively debated.

You correctly state that the University Committee reached opposing conclusions on the facts from Mr Justice Bokhary, who had tried the initial case, and you criticize them for doing so. You state "Hong Kong's Appeal Court confirmed an earlier court decision in favour of Dr Linda Koo", but there is a legal rule that a court of appeal, provided there is some evidence to justify a judge's factual findings, will not alter them at all (with rare exceptions). Consequently the factual decision in almost all civil cases — which are tried by a judge alone — depends only on one person's judgement. Such a legal system is obviously fallible, but at the moment we do not have a better one.

The important legal issue that Lam's case highlights is how a genuine review can be made in legal appeals of the truth or falsity of the facts found by a court without re-hearing all the evidence in every case. Fortunately for him, in Lam's case the members of the university committee (unlike the court of appeal) had a statutory personal responsibility to decide for themselves what were the facts, so that they could not just accept the findings

made by the judge without themselves hearing evidence.

The committee consisted of six distinguished professors. They received far more evidence than was called at the civil trial; this included the two research assistants of the protagonists, who each (contrary to the sworn evidence of Koo at the trial that she had kept her questionnaire locked away and had shown it to no one else) revealed that each of them at various times had been given a copy of her questionnaire, and that this had also been used in interviewing patients. They point out in their report (contrary to what you state) that Koo and Ho "decided not to appear before the Committee to assist in its investigation...however we did receive two written submissions" from them.

Perhaps more important for the future in this branch of science was the written evidence of some 26 epidemiologists, who had each seen and compared the two questionnaires. The committee in its report quoted one distinguished professor emeritus: "I am absolutely astonished at the claim that one is copied from the other...it is hard to imagine two more dissimilar questionnaires. I found the suggestion that one was copied from the other quite absurd."

The committee went on to say: "Similar views were expressed by the rest — that the questionnaires were quite different, that the Koo questionnaire was inferior to the Lam questionnaire, and that there was no evidence of one having been copied from the other." The committee's conclusion (on the entirety of the evidence which far exceeded that mentioned above) was that "The Committee finds that Dr. Lam's questionnaire was not copied from the Koo questionnaire."

You are wrong in stating that the committee "received written opinions on the similarity of Koo and Lam's questionnaires from 26 eminent epidemiologists": every single one said they were dissimilar.

The committee then examined the evidence of these distinguished scientists on the question whether instruments such as questionnaires (whatever the legal position) were in academic circles known to (or should), then or today, attract copyright or confidentiality after publication of the results derived from their use. (Koo had made her first public presentation of her results before Lam began to design his questionnaire, and had voluntarily given him her paper before publication.) This question was relevant to whether, if Lam had made use of parts of Koo's questionnaire (which they emphatically concluded he had not) such use would have been "disgraceful or dishonourable" conduct for disciplinary purposes.

The committee pointed out that:

"The reasons given for the free (i.e. uncopyrighted) use of questionnaires are to do with the nature of the scientific enterprise. Knowledge is cumulative, building on past efforts. The spirit and practice of free exchange of ideas increases progress in scientific discovery and knowledge. Feedback from colleagues, which plays an important role in clarifying one's ideas and in the development of research, becomes difficult in the absence of such exchanges. It is crucial, in order to test the validity of claims of discovery, that the methods of research undertaken by the person who claims to have made that discovery be replicated in different contexts. Verification and falsification, which are essential methods for validation and progress, require the free use of questionnaires and other methods of research. It was pointed out to us that free exchange was particularly important in medical research since the whole aim of such research is to promote the well being of the community."

The legal decision of Mr Justice Bokhary and the court of appeal establishes for better or for worse, unless and until some higher court pronounces otherwise, that instruments such as questionnaires are capable of attracting copyright and confidentiality, so that they may be used or ideas taken from them in effect only with the express permission of the author. But what, I ask, is the effect of this on the research process, as well as on validation of research? Does the scientific community accept that this should be so? And if not, what can or should be done?

John Griffiths

*Des Voeux Chambers,
10/F Bank of East Asia Building,
10 Des Voeux Road Central, Hong Kong*

SIR — Plagiarism, although universally condemned, is only occasionally punished, and perhaps the most recent supposed example shows why. Here and there a degree has been withdrawn or a resignation forced because of frank copying of the results of someone else's scholarly labours.

That is not the case with Dr Lam, so your harsh comments about him and the university are unjustified. If he deliberately copied anything at all, it was not the results of an investigation but certain of the tools (somewhat like Watson and Crick copying from Franklin and Wilkins?). How far must one go, exactly, in acknowledging the primacy of others when designing a questionnaire?

The university, investigating a matter of great weight for the career of a member of staff, was not compelled to adhere to the findings of a civil court on another question, decided on a balance of probabilities and on different facts.

T. R. C. Boyde

*123 Gloucester Place,
London W1H 3PJ, UK*

The hidden science of eugenics

Diane B. Paul and Hamish G. Spencer

The early eugenicists were not stupid, but they did not share our social values. The rise and fall of the eugenics movement is a history that modern medical geneticists would do well to heed.

MANY textbooks suggest that eugenicists were guilty of an astoundingly simple mistake. According to conventional accounts, enthusiasts about eugenics thought they could eliminate mental deficiency by segregating or sterilizing affected individuals. But a basic understanding of the Hardy–Weinberg principle suffices to destroy that illusion.

Eugenicists in the 1910s and 1920s attributed most mental defect to a recessive Mendelian factor (or in today's parlance, allele). But it is clear from the equation $p^2 + 2pq + q^2 = 1$ that if a trait is rare, most deleterious genes will be hidden in apparently normal carriers. Selection against those actually affected will thus be ineffective. Tables and formulas in many general biology and genetics textbooks (for example, refs 1–4) serve to make the point that hundreds of generations are required before a rare deleterious trait would disappear.

It is true that many eugenicists were muddled about genetics. But what about the host of respected geneticists, such as R. A. Fisher in the United Kingdom, Erwin Bauer in Germany, Herman Nilsson-Ehle in Sweden and Edward Murray East in the United States, who championed eugenics long after the implications of the Hardy–Weinberg principle were understood? The insight that selection is slow when genes are rare originated in 1917 and was popularized in the 1920s by J. B. S. Haldane in the United Kingdom and H. S. Jennings in the United States. Yet in the 1920s and 1930s, nearly all geneticists, including those traditionally characterized as opponents of eugenics, took it for granted that 'mental defectives' should be prevented from breeding. Moreover, the geneticists who first discussed the social implications of the Hardy–Weinberg principle did so in an effort to expand the scope of eugenics, not demonstrate its futility.

Invisible danger

In his 1917 essay "Hidden Feeble-mindedness"⁵, the Harvard geneticist East lauded efforts to cut off the stream of "defective germplasm" through segregation or sterilization of the affected. But East thought that the primary danger lay elsewhere, in the vast mass of invisible heterozygotes.

East had been strongly influenced by the psychologist Henry H. Goddard, author of *The Kallikak Family: A Study in the Heredity of Feeble-Mindedness*, a chron-

icle based on data collected by a field-worker who traced family members and assessed their mental and moral state. Many family members were of course dead or could not be located. Their mentality and character were assessed on the basis of hearsay. Judgements of both the living and dead were swift and subjective.

Two years later, in 1914, Goddard published *Feeble-Mindedness: Its Causes and Consequences*, in which he discussed the meaning of the Kallikak data for theories of inheritance. He argued that "normal-mindedness" is a dominant trait inherited in a Mendelian fashion; an individual lacking the factor for normal mentality would be feeble-minded — "incapable of performing his duties as a member of society in the position of life to which he is born"⁶.

The recessive theory of mental defect was widely accepted by Mendelians. The Cambridge geneticist R. C. Punnett spoke for many when he wrote that no one "who has studied the numerous pedigrees collected by Goddard and others [could] fail to draw the conclusion that this mental state [that is, feeble-mindedness] behaves as a simple recessive to the normal"⁷. Charles Davenport did note the illogicality of expecting a socially defined trait to be inherited as a simple Mendelian recessive; he thought there were many different (and separately inherited) mental deficiencies⁸. A few geneticists also contested Goddard's claim that feeble-mindedness was caused by a single Mendelian factor⁹. But these were minor quarrels. With the exception of Thomas Hunt Morgan, who argued that much of the behaviour associated with feeble-mindedness arose from "demoralizing social conditions"¹⁰, no Mendelian geneticist before the 1930s rejected Goddard's claim that social deviance was largely due to bad — recessive — heredity¹¹.

In 1912 Davenport offered the following advice:

Prevent the feeble-minded, drunkards, paupers, sex offenders, and criminalistic from marrying their like or cousins or any person belonging to a neuropathic strain. Practically it might be well to segregate such persons during the reproductive period for one generation. Then the crop of defectives will be reduced to practically nothing⁸.

At the time, such predictions were common. Even without the benefit of Hardy–Weinberg, East realized that they were wrong. The "real menace" of the

feeble-minded, he argued, lay in the huge heterozygotic reserve, constituting about seven per cent of the US population, or one in every fourteen individuals. East sounded an alarm: "Our modern Red Cross Knights have glimpsed but the face of the dragon."⁵

A question of time

His point was echoed by Punnett. For his influential *Mimicry in Butterflies* (1915), Punnett needed to know how fast a Mendelian factor would spread through a population¹². His Cambridge mathematics colleague, H. T. J. Norton, prepared a table displaying the number of generations required to change the frequency of completely dominant or recessive factors at different selection intensities. Punnett called attention to the table's implications for eugenics. Policies aimed at the affected, he argued, would take a distressingly long time to work. The Hardy–Weinberg formula indicated that more than ten per cent of the population carried the gene for feeble-mindedness. With G. H. Hardy's help, he also estimated the rate at which a population could be freed from mental defects by segregating or sterilizing the affected. Even under the unrealistic assumption that all the feeble-minded could be prevented from breeding, it would take more than 8,000 years before their numbers were reduced to 1 in 100,000, given Goddard's estimate that about 3 in 1,000 Americans were genetically feeble-minded. Punnett concluded that eugenic segregation did not, contrary to his initial belief, seem hopeful¹³.

Punnett, who served with Fisher on the Council of the Cambridge University Eugenics Society, did not intend to provide an argument against eugenics. In fact, he explicitly endorsed both East's scientific point and his policy proposals. Like East, he aimed to convince his readers of the need to identify the carriers of defective genes. "Clearly if that most desirable goal of a world rid of the feeble-minded is to be reached in a reasonable time," he asserted, "some method other than that of the elimination of the feeble-minded themselves must eventually be found."¹³

According to Fisher, Punnett's goal was subverted by opponents of eugenics, who seized on his table to argue that segregation and sterilization worked too slowly to justify the effort. In a 1924 article, "The Elimination of Mental Defect", Fisher argued that Punnett's calculations

obscured the fact that selection would initially be rapid. And for all practical purposes, he noted, that is what matters. Even under Punnett's unrealistic assumptions of a single gene for mental defect and of random mating, he argued, substantial progress could be achieved in the first few generations if affected individuals were prevented from breeding. In the first generation alone, the reduction would be more than 11 per cent¹⁴.

Notwithstanding some technical differences, Jennings, Punnett and Fisher agreed that mental defectives should be prevented from breeding. Jennings, who is sometimes portrayed as an opponent of eugenics, asserted that a gene that produces feeble-mindedness "is the embodiment, the material realization of a demon of evil; a living self-perpetuating creature, invisible, impalpable, that blasts the human being in bud or leaf. Such a thing must be stopped wherever it is recognized."¹⁵ Fisher diverged from Punnett and Jennings (who asserted as late as 1930 that feeble-mindedness was "the clearest case" of a recessive single gene defect¹⁶) only in claiming that the affected tended to mate with each other and that the trait was multifactorial. All agreed that the incidence of mental defect could be reduced by at least 10 per cent in the first generation and by 19 and 26 per cent by the second and third generations respectively. Even Haldane, who regarded compulsory sterilization "as a piece of crude Americanism", thought it "would probably cut down the supply of mental defectives in the next generation by something of the order of 10 percent"¹⁷.

Why were the estimates so high? It is often said that eugenics was based on a mistake about the efficacy of selection against rare genes. But few geneticists made this error, at least after 1917. Feeble-mindedness was not considered to be rare. Indeed, the *raison d'être* of the eugenics movement was the perceived threat of swamping by a large and rapidly growing class of mental defectives.

According to eugenicists, the number of physical and mental defectives would once have been kept in check by natural selection. But in civilized societies, it seemed that the process had practically ceased. Medicine and public charity now kept the 'unfit' alive. Worse, these failures were now reproducing faster than their betters. In the first three decades of this century, a raft of studies seemed to demonstrate the high fertility of the feeble-minded. For example, the British Royal Commission on the Care and Control of the Feeble-Minded reported in

1908 that defectives averaged seven children, normal couples only four. In the United States it was commonly believed that between 300,000 and 1,000,000 people were feeble-minded as a result of genetic defects⁹; the figures tended to increase as mental tests came into wider use. In 1912, Goddard tested New York City schoolchildren and estimated that two per cent were probably feeble-minded. The results of tests given to army recruits during the First World War were even

other diseases²⁰. Selection here is futile.

But these arguments do not apply to the early eugenics movements. Given widely shared assumptions about the incidence and causes of feeble-mindedness, eugenic policies could be expected to reduce substantially the number of affected individuals. In the famous case of *Buck v. Bell* (1927), decided in the wake of US Army mental tests, Justice Holmes upheld the Virginia sterilization law "to prevent our being swamped with incompetence". Commenting on this passage, one author remarks that "what such reasoning fails to take into account is that. . . sterilization will have almost no effect on the frequency of the disease"²¹. The point is illustrated by a recessive disorder affecting 1 in 40,000 individuals.

In any case, geneticists in the 1920s would generally have favoured eugenic policies whatever their exact effect. Most would have assented to Jennings's claim that "to stop the propagation of the feeble-minded, by thoroughly effective measures, is a procedure for the welfare of future generations that should be supported by all enlightened persons. Even though it may get rid of but a small proportion of the defective genes, every case saved is a gain, is worthwhile in itself"¹⁶.

Like Jennings, Lancelot Hogben is often portrayed as an opponent of eugenics. Hogben did criticize some advocates of sterilization for exaggerating both the extent of the problem and efficacy of their solution. He also argued that the fact that we cannot do everything "is not a valid reason for neglecting to do what little can be done"²². His point was echoed by E. G. Conklin, who, like Jennings and Hogben, criticized some eugenic proposals. But Conklin approved of the segregation or sterilization of the feeble-minded. He asked of the American Eugenics Society's proposed sterilization policy: "Can any serious objection be urged to such a law?"²².

Punnett, East, Fisher, Jennings and even Haldane made roughly the same estimates of the speed and scope of eugenic selection. But the facts did not speak for themselves. They required interpretation in the light of other assumptions and goals. Thus Haldane opposed sterilization, arguing that "with mental defects as with physical defects, if you once deem it desirable to sterilize I think it is a little difficult to know where you are to stop"¹⁷. This is a powerful argument. But it is a social, not a scientific one. Lionel Penrose was the most vehement geneticist critic of eugenics. An expert in the genetics of mental deficiency, he stressed the heterogeneity of its causes and the modest influence of eugenic measures in reducing its incidence. But his main argument was ethical. Penrose maintained that the best index of a society's health is its willingness to provide adequate care for those unable

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Breeding propaganda — London Eugenics Society poster, c. 1935.

more dramatic: they indicated that 47.3 per cent of the white draft and 89 per cent of the black draft were feeble-minded¹⁸.

Contemporary textbook examples of the futility of eugenics often mention Tay-Sachs disease, phenylketonuria and albinism. Selection against such conditions is certainly futile. But they are usually rare and their effects either lethal or minor. Victims of Tay-Sachs and (with a few exceptions) untreated phenylketonuria do not leave offspring. Albinos and treated phenylketonurics reproduce, but these conditions are not disabling. And with regard to the potential increase in genes for treatable genetic diseases, the Hardy-Weinberg argument is relevant¹⁹. Most single-gene disorders are extremely rare and would spread slowly in the population. Moreover, some disease genes may be maintained at high frequency because of heterozygote advantage; although this is difficult to demonstrate, some geneticists believe it explains not only the well-confirmed case of sickle-cell anaemia, but also cystic fibrosis and Tay-Sachs (in Ashkenazi Jews), among

Wellcome Institute Library

to care for themselves²³.

The Hardy-Weinberg theorem meant different things to different people. For example, Curt Stern once remarked: "To state that reproductive selection against severe physical and mental abnormalities will reduce the number of affected from one generation to the next by only a few per cent does not alter the fact that these few per cent may mean tens of thousands of unfortunate individuals who, if never born, will be saved untold sorrow"²⁴. It may not even matter if the reduction in absolute numbers is minuscule; the rate of selection is immaterial if one assumes with Jennings that the "prevention of propagation of even one congenitally defective individual puts a period to at least one line of operation of this devil. To fail to do at least so much would be a crime."²⁵

Many advocates of sterilization employed a loose definition of 'feeble-mindedness', accepted Goddard's defective data and logic or assumed that "it would be possible at one fell stroke [to] cut off practically all of the cacogenic varieties of the race"²⁵. But it was possible to recognize all these flaws and still remain a eugenicist. After 1920, it was well understood that most genes for mental defects would be hidden in apparently normal carriers. For most geneticists this seemed to be a good reason to widen eugenic efforts rather than abandon them.

This implication makes sense in the light of social values. In 1918, Popenoe and Johnson wrote that "so few people would now contend that two feeble-minded or epileptic persons have any 'right' to marry and perpetuate their kind, that it is hardly worth while to argue the point"⁹.

Politics changed these assumptions. During the 1940s and 1950s, many geneticists tried to distinguish the race- and

class-biased eugenics of the past from a new eugenics that focused on disease. But attempts to distinguish good from bad eugenics were ultimately unsuccessful. Nazi atrocities gave eugenics of any kind a bad name and produced a backlash against the view that the state had a legitimate interest in who reproduced.

That reproduction should be a private matter was strongly reinforced by a trend towards respect for patients' medical rights, the development of a broad jurisprudence of privacy and the rise of feminism. By the 1960s, reproductive autonomy had become a dominant cultural value. This was a far cry from the assertion of a 1914 committee of the American Breeders Association: "Society must look upon germ-plasm as belonging to society and not solely to the individual who carries it"²⁵. A change in values, and not the progress of science, explains why few Swedes would now agree with the 1936 commission that criticized as "extremely individualistic" the notion that individuals have a right to control their own bodies²⁶.

It is often said that support for eugenics declined in the 1930s as its scientific errors were exposed. But the eugenics movement grew stronger during the Depression. In the United States, the number of sterilizations increased. Sterilization was legalized in Germany (1933), British Columbia, Canada (1933), Norway (1934), Sweden (1934), Finland (1935), Estonia (1936) and Iceland (1938). Denmark, which in 1929 had legalized 'voluntary' sterilization, permitted its coercive use on mental defectives in 1934. These laws were generally applauded by geneticists.

How then do we account for the popularity of the claim that eugenics was based on a technical error? We suggest two related reasons. In the first quarter of this

century, nearly all geneticists were enthusiastic proponents of a movement that is now generally held in contempt. In Germany, not one geneticist criticized the interwar eugenics movements²⁷. After the Nazis came to power, genetics was invoked on behalf of ever more extreme measures of racial purification. Nevertheless, most of Germany's leading geneticists, including those who before 1933 had criticized antisemitism, actively helped to build the racial state. They served on important commissions, provided opinions on racial ancestry and participated in the drafting of racial laws. More than half of all academic biologists joined the Nazi Party, the highest membership rate of any professional group²⁸. In other countries, too, eugenicists promoted policies such as immigration restriction that reflected strong class and racial biases. So the history of the field is the source of some embarrassment (and defensiveness). It is far more comforting to think that eugenics' decline was also due to geneticists. The myth rights the historical balance.

Backdoor eugenics

The claim also enables textbook writers and college teachers to avoid controversial issues. If eugenics is assumed to have rested on a technical error, it no longer raises thorny ethical questions. Geneticists can therefore condemn eugenics without questioning any of the aims of genetic testing. As Arthur Caplan points out, when the State of California ruled that screening for maternal serum α -feto-protein should be offered to all pregnant women, it did so "in the hope that some of those who are found to have children with neural tube defects will choose not to bring them to term. . . thereby preventing the state from having to bear the burden of their care"²⁹. There is similar cost-benefit reasoning in the 1990 guidelines of the International Huntington Association and the World Federation of Neurology, which deem it acceptable to refuse to test women who "do not give complete assurance that they will terminate a pregnancy where there is an increased risk" of Huntington's disease³⁰. Those who made this recommendation certainly did not think they were promoting eugenics. Assuming that eugenics is dead is one way to dispose of deep social, political and ethical questions. But it may not be the best one. □

Diane B. Paul is in the Department of Political Science, University of Massachusetts, Boston, Massachusetts 02125, USA; Hamish G. Spencer is in the Department of Zoology, University of Otago, PO Box 56, Dunedin, New Zealand.

ACKNOWLEDGEMENTS. We thank the Division of Sciences, University of Otago, for financial support; E. F. Keller, R. C. Lewontin, J. Maynard Smith, M. Ruse, E. Seaman and G. P. Wallis for reading the manuscript; L. Zenderland for comments on Goddard; and the staff of the Science Library, University of Otago, for help with references.

- Cummings, M. R. *Human Heredity: Principles and Issues* 2nd edn, 436-437 (West, St Paul, 1991).
- Curtis, H. & Barnes, N. S. *Biology* 5th edn, 987 (Worth, New York, 1989).
- Griffiths, A. J. F., Miller, J. H., Suzuki, D. T., Lewontin, R. C. & Gelbart, W. M. *An Introduction to Genetic Analysis* 5th edn, 759-760 (W. H. Freeman, New York, 1993).
- Strickberger, M. W. *Genetics* 3rd edn, 693 (Macmillan, New York, 1985).
- East, E. M. *J. Hered.* **8**, 215-217 (1917).
- Popenoe, P. *J. Hered.* **6**, 32-36 (1915).
- Punnett, R. C. *The Nineteenth Century* **97**, 697-707 (1925).
- Davenport, C. B. in *Heredity and Eugenics* (eds Castle, W. E., Coulter, J. M., Davenport, C. B., East, E. M. & Porter, W. L.) 286 (Univ. Chicago Press, Chicago, 1912).
- Popenoe, P. & Johnson, R. H. *Applied Eugenics* 105, 170-176 (Macmillan, New York, 1918).
- Morgan, T. H. *Evolution and Genetics* 2nd edn, 201 (Princeton Univ. Press, Princeton, 1925).
- Barker, D. *Brit. J. Hist. Sci.* **22**, 347-375 (1989).
- Provine, W. B. *The Origins of Theoretical Population Genetics* 137 (Univ. Chicago Press, Chicago, 1971).
- Punnett, R. C. *J. Hered.* **8**, 464-465 (1917).
- Fisher, R. A. *Eugenics Rev.* **26**, 114-116 (1924). (Reprinted in *Collected Papers of R. A. Fisher* (ed. Bennett, J. H.) Vol. 1, 580-582 (Univ. Adelaide, South Australia, 1971).)
- Jennings, H. S. *J. Hered.* **18**, 271-276 (1927).
- Jennings, H. S. *The Biological Basis of Human Nature* 238-242 (Faber & Faber, London, 1930).
- Haldane, J. B. S. *Heredity and Politics* 80-89 (Unwin, London, 1938).
- Brigham, C. C. *A Study of American Intelligence* (Princeton Univ. Press, Princeton, 1923).

- Paul, D. B. & Spencer, H. G. *Am. J. hum. Genet.* **43**, 344-346 (1988).
- Emery, A. E. H. & Mueller, R. F. *Elements of Medical Genetics* 8th edn (Churchill Livingstone, Edinburgh, 1992).
- Hogben, L. *Genetic Principles in Medicine and Social Science* 207 (Williams & Norgate, London, 1931).
- Conklin, E. G. in *Human Biology and Population Improvement* (ed. Cowdry, E. V.) 577-578 (Hoebner, New York, 1930).
- Penrose, L. S. *The Biology of Mental Defect* (Sidgwick & Jackson, London, 1949).
- Stern, C. *Principles of Human Genetics* 538 (W. H. Freeman, San Francisco, 1949).
- Laughlin, H. H. *Report of the Committee to Study and to Report on the Best Practical Means of Cutting off the Defective Germ-Plasm in the American Population — Vol. I: The Scope of the Committee's Work* 47, 16 (Eugenics Record Office, Bulletin No. 10A, Cold Spring Harbor, New York, 1914).
- Broberg, G. in *Eugenics and the Welfare State: Sterilization Policy in Denmark, Sweden, Norway, and Finland* (eds Broberg, G. & Roll-Hansen, N.) (Michigan State Univ. Press, East Lansing, in the press).
- Harwood, J. *Brit. J. Hist. Sci.* **22**, 257-265 (1989).
- Deichmann, U. & Müller-Hill, B. in *Science, Technology and National Socialism* (eds Renneberg, M. & Walker, M.) 160-183 (CUP, Cambridge, 1994).
- Caplan, A. L. in *Prescribing Our Future: Ethical Challenges in Genetic Counselling* (ed. Bartels, D.) 149-165 (Aldine de Gruyter, Hawthorne, New York, 1993).
- Ethical Issues Policy Statement on Huntington's Disease Molecular Genetics Predictive Test *J. Med. Genet.* **27**, 34-38 (1990).

Sustainable development unsustainable

Environmentalists tempted to ignore Wilfred Beckerman's latest book had better think again; there is no better way of 'knowing your enemy', and no more mordant enemy.

YET another book review preceded by the disclaimer that the author is a friend or, in this case, worse — an ally! The author is the Balliol (Oxford) economist Wilfred Beckerman. The title, an obvious parody of that of E.F. ("Fritz") Schumacher's *Small is Beautiful* (Blond, 1973), is *Small is Stupid* (Duckworth, 1995). But Beckerman, whose professional field of expertise is the national accounts of nation-states such as Britain, has not on this occasion written an economic textbook, but a tract. That, in a sense, is what Schumacher did, but in gentler language.

The circumstances of my alliance with Beckerman are these. In 1971, he was a member of the British Royal Commission (standing committee) on Environmental Pollution and in that role was bombarded with photocopies of articles appearing in the press at large, *Nature* included, on environmental matters. (The commission's otherwise excellent secretariat seems to pay no attention to copyright law.) At some stage, Beckerman wrote to say that some article in *Nature* seemed to him to take a level-headed view of the economics of pollution, so when it came to debating an environmental topic at the Oxford Union, I recruited him as my second. (He was excellent, but we lost.)

It was there I first heard his joke (retold in this book) about the element called 'Beckermanium', allegedly named after his grandfather who had failed to discover it. Has the world been the worse for his neglect is the rhetorical punchline, an oblique comment on the assertion that the industrial economies of the world are quickly using up the minerals on which they apparently depend.

The joke apart, environmentalists will not find the book a comfortable read. The subtitle, "Blowing the whistle on the Greens", is a sufficient warning. And Beckerman does enjoy making fun of false past predictions of catastrophe; the table in which he compares the reserves of metals such as copper and lead quoted in *The Limits to Growth* (1972) with the corresponding values for 1989 is headed, with studied inelegance, "How we used up all the resources we had and still finished up with more than we started with".

To take the metal lead as an example, consumption between 1970 and 1989 amounted to 99 million tonnes, rather more than the 91 million metric tonnes of reserves supposed to exist in 1970, yet by 1989 the known reserves of lead had actu-

ally increased, to 125 million tonnes. There, Beckerman might have said, is a widow's cruse for you! This and similar arguments point him to the conclusion that there must be "major flaws" in the argument that the Earth's resources are finite; surprisingly, for an economist, he does not suggest that what matters in the real world is the distribution of reserves with price (or extraction cost).

But environmentalists rather than those already converted to Beckerman's view are the ones who should read his book. Its purpose is to take apart the now-familiar notions of "sustainable development" and "precautionary principle". The former appears in the Brundtland Report (named in 1987 after its woman chairman, the prime minister of Norway, and entitled *Our Common Future*). For on the principle that one should know one's enemy, environmentalists need to know why some say that "sustainable development" is meaningless in the strict sense that it is empty of operational meaning. "The Emperor of Sustainable Development has no clothes", says Beckerman.

He is surely right. It is not only a matter of definition, although the definitions matter. Brundtland's injunction, for example, is that sustainable development must meet "the needs of present generations without compromising the ability of future generations to meet their own needs". It is not a frivolous question to ask how it can be possible with any certainty to know what future needs will be.

In the absence of more information, it is reasonable to suppose that future needs and present needs will be identical, but the logical consequence is that it would be improper of the present generation to make any use of a nonrenewable resource, say lead. To do so would mean that future generations would be less free than we are to mine lead, but on the assumption that the goal of sustainable development is itself sustainable across the generations, future generations would also be interdicted (by Brundtland) from using lead. Nonrenewable resources would be untouchable unless one generation could persuade itself that its successors will have less need of lead than we have.

There are many contrary examples in economic history. In South Wales in the 1920s, for example, the output of coal from the coalmines amounted to 250 million tons a year, of which 100 million tons were exported from Welsh ports to coun-

tries outside Britain. Evidently the notion of sustainable development had not caught on. Even if the extraction of the domestic demand for coal (150 million tons a year) could be justified on the grounds that there seemed to be plenty more where that came from, physically exporting 100 million tons a year of it must surely have 'compromised' the ability of future (that is, present) generations to 'meet their own needs'.

The snag, of course, is that coal consumption has slumped. Deep-mined coal production from the region is now less than 1 per cent of what it was 70 years ago. So in what way would it have benefited South Wales, either in the 1920s or now, if it had been resolved not to extract coal in excess of local demand? The region would have been less prosperous than it was and thus less able to build the excellent school system that was, for many years, its hallmark. Now there is no coalmining industry worth speaking of.

Beckerman, in his book, argues that the sustainability engineers have been watering down sustainability ever since Brundtland appeared. He makes easy game of his compatriot-economist David Pearce, who most recently subscribes to a definition in terms of wellbeing (which must not deteriorate over the generations), but which allows the substitution of man-made for 'natural' capital on the surface of the Earth. Is not that just another way of saying the 'welfare' (of which the predominant component is gross national product) is all that matters?

The 'precautionary principle', the doctrine that one should act now in case things go wrong later, unravels in much the same way, but Beckerman the economist is more obviously in charge. The extra cost of acting too soon is never calculated accurately. And putting the precautionary principle into effect often destroys the evidence of its necessity.

In short, Beckerman has written a thoughtful book as well as an entertaining one. Environmentalists will not be too badly hurt by the jibes. But Beckerman is probably wrong in putting global warming in the same bag as the other green excesses. He is right, of course, to insist that the cost of delaying action would be less than the global warmers say, but the interval should be used as a time of preparation to take action. That, luckily, is what this month's meeting in Berlin is all about.

John Maddox

Zippering together a cell adhesion interface

Dinshaw J. Patel and Barry M. Gumbiner

CADHERINS are molecules that are responsible for selective cell recognition and adhesion during the morphogenesis of the embryo, and for the maintenance of normal tissue architecture throughout the life of an adult organism¹. Clearly they are of immense biological interest, and a molecular view of how they work emerges in the paper by Shapiro *et al.* on page 327 of this issue².

The authors provide a single-crystal, X-ray characterization of the amino-terminal domain of murine N (for neural) cadherin (NCD1), along with analysis of five independent structures available in

crystal forms with Yb^{3+} replacing Ca^{2+} , and in a third form with UO_2^{2+} in the Ca^{2+} coordination site. The data were collected using the new multiwavelength anomalous diffraction (MAD) approach, and established for the first time the feasibility of combining undulator radiation and MAD phasing techniques in the case of the NCD1- Yb^{3+} complex. The structure has been solved to high resolution (1.9 to 2.1 Å). It turns out that the polypeptide fold of the protomer consists of a seven-stranded β -sandwich, supplemented by an unusual quasi β -helix and a short 3_{10} helix. The multivalent cation-binding site is

A second dimer interface, also with approximate two-fold symmetry and common to all crystal forms, involves the antiparallel alignment of protomers from opposing cell surfaces. Dimerization at this adhesion dimer interface involves an even larger surface area (3,300 Å²) composed of residues from β -strands C, D, F and G and the quasi β -helix (Fig. 1, right). This large interface includes the HAV segment on the F strand which has been implicated in homophilic cadherin interactions³. The authors suggest that the adhesive dimer interface is weaker than the strand dimer interface, because the first is characterized by water-mediated and long-range interactions whereas the second is associated with close-packed hydrophobic interactions.

This emphasis on cadherin dimers as the fundamental unit for formation of the adhesive interface contrasts with an independent NMR-based structure of the corresponding domain of E (for epithelial) cadherin, where the structure analysis was based on a monomeric species in solution⁴. However, gel-filtration and dynamic light-scattering experiments indicate that an NCD1 dimer forms in solution², and Shapiro and colleagues propose that this corresponds to the strand dimer.

The Yb^{3+} and UO_2^{2+} -coordinated NCD1 structures suggest that Ca^{2+} ions bridge successive domains through coordination of acidic residues on adjacent protomers. The proposed positioning of adjacent strand dimers in an end-on-end alignment with Ca^{2+} coordination requires a 120° rotation about the strand dimer axis. This proposed alignment is testable by extending the crystal-structure analysis to an NCD1-NCD2 construct, which should define not only the coordination of the Ca^{2+} -binding site but also the role of the linker segment in determining the relative alignment and interaction between sequential domains. The central involvement of Ca^{2+} in orienting and stabilizing successive cadherin domains in the model is supported by experimental data⁵, which show that adhesive interaction between cadherins maintains its integrity only in the presence of this divalent cation.

The proposed cell-adhesion zipper has been constructed by modelling the interdigitative interaction of strand and adhesion dimers, along with the Ca^{2+} -coordinated sequential linkage of successive domains. First, the combinatorial alignment of strand dimers and adhesion dimers generates a continuous adhesive ribbon (Fig. 2), which is a common feature of all three crystal forms. The molecular basis for cell adhesion resides in this ribbon and involves NCD1 dimers projecting from opposing cell surfaces. Second, the association of Ca^{2+} -coordinated sequential domains bridges the space between

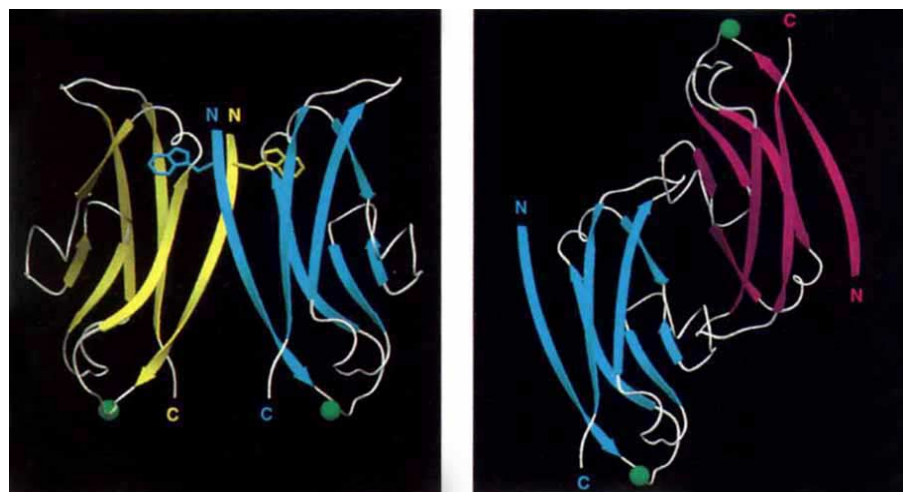


FIG. 1 Left, Ribbon drawing of the strand dimer with individual protomers in cyan and yellow. The termini are denoted by N and C, and uranium atoms are shown as green spheres. The side chain of Trp 2 from each protomer extends into a hydrophobic pocket of its partner in the dimer. Right, Ribbon drawing of the adhesion dimer with individual protomers in cyan and magenta.

three separate crystal forms. The key element was the identification and functional assignment of specific dimer interfaces, and the subsequent construction through interdigitative alignments of an adhesive superstructure. This structure greatly contributes to our understanding of the mechanism of cadherin-mediated adhesion between cells.

The cadherins are single-pass transmembrane glycoproteins that mediate homophilic adhesion between neighbouring cells. Their extracellular domains are involved in Ca^{2+} -dependent cell-cell adhesion; their cytoplasmic domains mediate connections with cytoskeletal actin filaments in *zonula adherens* junctions and with intermediate filaments in desmosome junctions. The extracellular domain generally consists of five tandem repeat sequences, of some 110 amino-acids each, with the adhesive specificity residing in the amino-terminal domain.

The NCD1 structure was solved in two

positioned between two loops towards the carboxy terminus of the protomer. The folding topology includes three *cis* prolines with the amino and carboxy termini positioned at opposite ends of the protomer module.

Shapiro *et al.* identified a twofold symmetrical dimer interface, involving parallel alignment of protomers from the same cell surface, in all three crystal forms, and designated it the strand dimer interface. Dimerization involves a large surface area (1,800 Å²) containing the highly conserved amino-terminal residues of β -strand A, and is characterized by the mutual insertion of the side chain of Trp 2 into a hydrophobic pocket of the partner protomer across the interface (Fig. 1, left). Such a conserved strand dimer interface between NCD1 domains is probably also adopted by NCD domains 2 to 5, contributing to lateral dimerization of the entire cadherin molecule in the extracellular region.

the adhesive ribbon and opposing cell surfaces, completing the proposed interdigitated lattice of cell–cell adhesion by cadherins. So weak monomeric associations spread through interdigitative interactions over an extensive lattice can result in a cooperative cell–cell recognition network of great strength.

The proposed zipper model for N-cadherin should be a prototype for cell–cell interactions involving other members of the family. This is because cadherins show a high degree of conservation of residues at the strand dimer interface, and have complementary pairs of species-specific side-chain substitutions of residues at the adhesion dimer interface. Such complementarity between residues across this interface is also maintained for heterophilic cell–cell adhesion between retinal and neural cadherins.

The folding topologies of the CD1 domains of N- and E-cadherins are similar to those reported for immunoglobulin-like molecules^{2,4,6}, raising the issue of whether the cadherins are distant members of the immunoglobulin superfamily⁷. An analysis of sequence comparisons and intron patterns, however, seems to favour the

view that the two originated independently, rather than diverged from a common ancestor⁶. Structures of specific cell–adhesion interfaces for immunoglobulins remain to be determined, and it will be of great interest to compare the intermolecular alignments and networks with the cell–adhesion zipper proposed for the cadherin superfamily.

Another outstanding matter is that of how the adhesive activities of cadherins at the cell surface are regulated by cytoplasmic signals⁸. A complete functional view of cadherins must invoke their transmembrane alignment and the association of their intracellular components with the cytoskeleton as mediated by catenins. The work of Shapiro *et al.* provides a framework not only for probing cadherin

structure in more detail, but also for tackling questions about the higher-order structure of adhesive junctions and the mechanisms that dynamically regulate cadherin-mediated cell adhesion. □

Dinshaw J. Patel and Barry M. Gumbiner are in the Cellular Biochemistry and Biophysics Program, Memorial Sloan-Kettering Cancer Center, New York, New York 10021, USA.

1. Takeichi, M. *Science* **251**, 1451–1455 (1991).
2. Shapiro, L. *et al.* *Nature* **374**, 327–337 (1995).
3. Blaschuk, O. W., Sullivan, R., David, S. & Pouliot, Y. *Dev Biol.* **139**, 227–229 (1990).
4. Overduin, M. *et al.* *Science* **267**, 386–389 (1995).
5. Pokutta, S., Herrenknecht, K., Kemler, R. & Engel, J. *Eur. J. Biochem.* **223**, 1019–1026 (1994).
6. Shapiro, L. *et al.* *Proc. natn. Acad. Sci. U.S.A.* (in the press).
7. Wagner, G. *Science* **267**, 342 (1995).
8. Gumbiner, B. M. *Neuron* **11**, 551–564 (1993).

NONLINEAR OPTICS

Propeller molecules take off

Graham Cross

'SECOND-order' nonlinear optical effects (such as second harmonic generation and the linear electro-optic effect) underpin many of the active functions required in optoelectronics, and new materials showing high second-order responses are constantly sought. Organics in particular have much to offer. But high molecular nonlinearity is usually tied to a high dipole moment, which can cause problems. Dhenaut and colleagues at CNET, writing on page 339 of this issue¹, have now explored the possibility that a large molecular second-order nonlinearity could co-exist with a complete net absence of dipole moment. In the ruthenium(II) trisubstituted bipyridine complex (otherwise known as a 'propeller' complex because of its geometry) the group seems to have achieved its aim.

In the field-induced polarization, p_{ind} , of a non-centrosymmetric molecule in an electric field E ($p_{ind} = \alpha E + \beta EE + \gamma EEE + \dots$) the coefficient β , which is known as the first hyperpolarizability, governs the magnitude of the second-order response. For comparison between materials, one must compare values of β between molecules in the absence of resonant enhancement (the term $\beta(0)$ is used to denote the zero-frequency value). The reported value for $\beta(0)$ of 10^{-27} e.s.u. for the new material compares very favourably with the current leading materials^{2–5}, all the rest of which possess large dipoles.

In the mainstream studies of molecular second-order nonlinearity, a straightforward design strategy has been followed since the mid-1970s. The model structure *p*-nitroaniline (*p*-NA; part *a* of the box) demonstrates the generic features of this strategy. In *p*-NA, a polarizable bridging group — in this case an aromatic group —

links functional groups of differing electron affinity. The electronic bias results in a static dipole moment and the second-order nonlinearity results from rectification in the induced polarization under the action of applied fields. Although β is in fact a third-rank tensor the rod-like geometry of the *p*-NA model compounds ensures that only a single axial component dominates.

Importantly, the molecule is non-centrosymmetric: inverting the molecule with respect to a fixed frame of reference produces a dipole of the opposite sign. This non-centrosymmetry must be carried through into bulk properties to achieve macroscopic second-order nonlinear optical effects. Molecules may be forced into alignment using, for example, the Langmuir–Blodgett technique⁶ or, alternatively, molecules with high dipole moments (perhaps dispersed into a polymer) may be aligned using electric fields. The presence of dipoles in the standard materials is thus seen as an essential feature. So the new non-dipolar molecules developed by Zyss's group at CNET^{1,7–9} represent an almost iconoclastic departure from conventional ideas.

In designing non-dipolar molecules with non-zero β , the group has looked at molecules with a higher degree of symmetry than the rod-like *p*-NA analogues. In initial work, they found hexa-substituted benzene, TATB (see box), a trigonal analogue of *p*-NA, to be a model compound⁷. Second-harmonic generation measurements on the powder showed that TATB must possess a finite β and that its crystal structure must have orientational order (the concept of 'polar order' being inappropriate for this class of materials). The molecule has planar trigonal

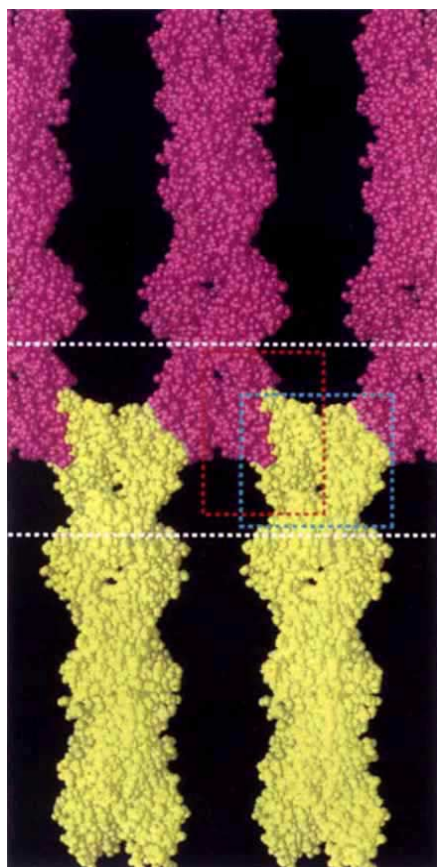


FIG. 2 Space-filling model of the proposed complete cell-adhesion zipper. Molecules in yellow and magenta correspond to cadherin extracellular segments emanating from opposing cell surfaces. The strand dimer is enclosed in a dashed cyan box, the adhesion dimer in a dashed red box and the cell-adhesion ribbon in a dashed white box.

symmetry with a three-fold rotation axis. All dipolar-like features (including the dipolar β responses) sum to zero in such a system. So how does a finite β remain? The answer is far from intuitive, but this observation in TATB, and in the new molecule reported in this issue, has been rationalized by Zyss⁸ in terms of the octupolar response of the molecule (see part *b* of the box).

By wrestling with group theory analysis Zyss showed that in general the β tensor may contain contributions from odd-order multipoles. Thus in molecules of particular symmetries (for example trigonal or tetrahedral) significant octupolar contributions to the β tensor can arise. The octupolar part of the β tensor has many more independent components than the dipolar part (which vanishes anyway in octupolar molecules). When taken collectively, these components effectively pack more nonlinearity into a single molecule. Thus the use of octupolar geometries may allow a larger molecular nonlinearity per unit volume; a nonlinearity less sensitive to material orientation (a higher dimensionality of nonlinearity exists); and all without the possible detrimental effects of dipole moment.

The ruthenium complex reported in this issue clearly benefits from the additivity principle and has chirality built in to ensure octupolar crystallinity. The value of the off-resonance octupolar β is indeed very large (in fact it exceeds the predicted tensor component sum by a factor of two)

and the question now is how to make use of it. Crystals may show the highest bulk nonlinearity, but their processability is limited; and Langmuir–Blodgett assembly is still not favoured by those looking for simple mass production. The preferred choice may be to incorporate the molecule into a polymer and perform some kind of ‘poling’. Usually, this means dragging dipolar molecules into alignment by applying an electric field. In the absence of a dipole, conventional poling methods are ruled out, but collaborators of the CNET group may have the solution¹⁰. Termed ‘orientational hole burning’, the method relies on selective absorption by, and thus effective bleaching of, just those molecules in a specific orientational range, preventing total cancellation of the bulk nonlinearity. The orientational distribution of molecular β s, now being asymmetric, allows for the bulk second-order effects.

Such an orientation scheme could be applied equally to polar molecules as to the new octupolar species. However, there must be question marks attached to the added complexity of the process and to the need for a resonant interaction with the material. Many of the more highly nonlinear molecules are also somewhat unstable to photo-induced degradation. The use of metal centres, as in the ruthenium complex, may ameliorate these problems, but by how much is unclear. The problem of ‘locking in’ the orientation is a further point to be addressed,

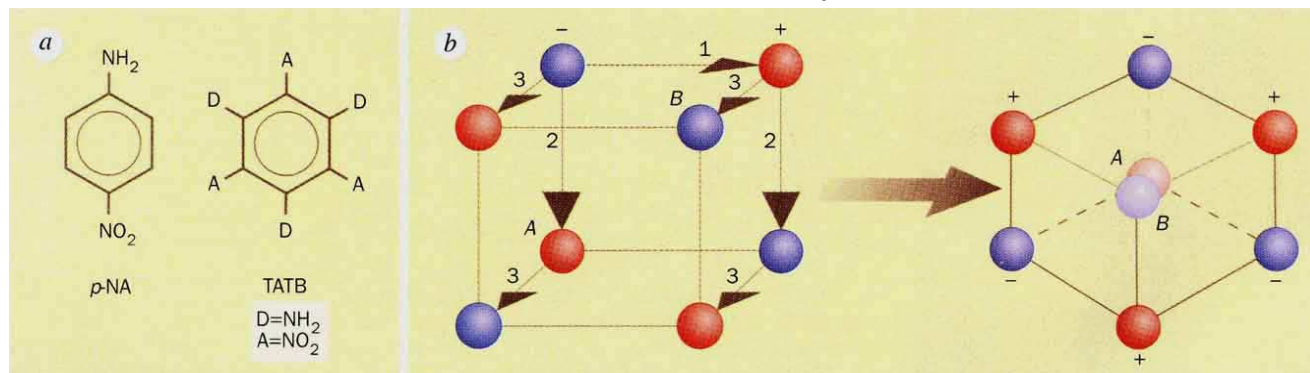
although there might be useful parallels in the observation of photo-induced excitation and *trans*-to-*cis* isomerization seen in azo-dyes^{11,12}.

The key advance from which these developments have grown rests on the connection that Zyss made between the rather abstract areas of tensor analysis and group theory as applied to optical nonlinearities and the realization that molecules could be designed which show octupolar nonlinearities. The potential advantages are self-evident. What remains to be seen is whether other workers in the field are willing, or able, to move up from dipoles to octupoles in their search for improved nonlinear materials. □

Graham Cross is in the Department of Physics, University of Durham, South Road, Durham DH1 3LE, UK.

1. Dhenaut, C., Ledoux, I., Samuel, I. D. W., Zyss, J., Bourgaud, M. & Le Bozec, H. *Nature* **374**, 339–342 (1995).
2. Marder, S. R. *et al.* *Science* **263**, 511–514 (1994).
3. Laidlaw, W. M., Denning, R. G., Verbiest, T., Chaudard, E. & Persoons, A. *Nature* **363**, 58–60 (1993).
4. Szablewski, M. J. *org. Chem.* **59**, 954–956 (1994).
5. Rao, V. P., Jen, A. K.-Y., Wong, K. Y. & Drost, K. J. *J. chem. Soc. chem. Commun.* 1118 (1993).
6. Peterson, I. R. in *Molecular Electronics* (ed. Ashwell, G. J.) 117–206 (Wiley, Chichester, 1992).
7. Ledoux, I., Zyss, J., Siegel, J. S. & Lehn, J.-M. *Chem. Phys. Lett.* **172**, 440–444 (1990).
8. Zyss, J. *Nonlinear Optics* **1**, 3–18 (1991).
9. Zyss, J., Dhenaut, C., Chauvan, T. & Ledoux, I. *Chem. Phys. Lett.* **206**, 409–414 (1993).
10. Nunzi, J.-M., Charra, F., Fiorini, C. & Zyss, J. *J. chem. Phys. Lett.* **219**, 349–354 (1994).
11. Jones, C. & Day, S. *Nature* **351**, 15 (1991).
12. Charra, F., Kajzar, F., Nunzi, J. M., Raimond, P. & Idiart, E. *Opt. Lett.* **18**, 941–943 (1993).

From dipoles to octupoles



a, *p*-NA is a model chromophore for second-order nonlinear optics. The general scheme is donor; π -conjugated group; acceptor. The single dominant component of β lies along the dipolar axis and its magnitude depends on the magnitude of $\Delta\mu$, the difference between ground and excited-state dipole moments. TATB is a trigonal symmetric version of *p*-NA having a complete net absence of dipole moment. The dipolar or vectorial part of the β tensor vanishes but the symmetry is not high enough to prevent the presence of a finite octupolar β tensor component. *b*, An octupole can be generated by successive translations and inversion of signs of point charges. Thus a dipole, more familiarly viewed as two equal and opposite charges separated by a certain distance, can be generated by a simple replication along a vector of a single point charge

followed by a reversal of the sign of the new charge. After a further two operations of this type a cube is generated. This has 2^n (n being the number of operations) or 8 point charges and is thus an octupole. To see how this three-dimensional object maps onto the two-dimensional TATB molecule, one must project the cube along one of the body diagonals, whereupon A and B coincide and cancel. The octupole is a non-dipolar yet non-centrosymmetric structure and thus satisfies the requirements for a non-zero β tensor. The full treatment of multipolar groups shows that structures possessing multipolar groups of higher order than $n = 3$ will not possess a finite β . Interestingly, and indeed non-intuitively, this applies to structures with five-fold symmetry ($n = 5$) which are, nevertheless, non-centrosymmetric. G.C.

The cheetah controversy

Robert M. May

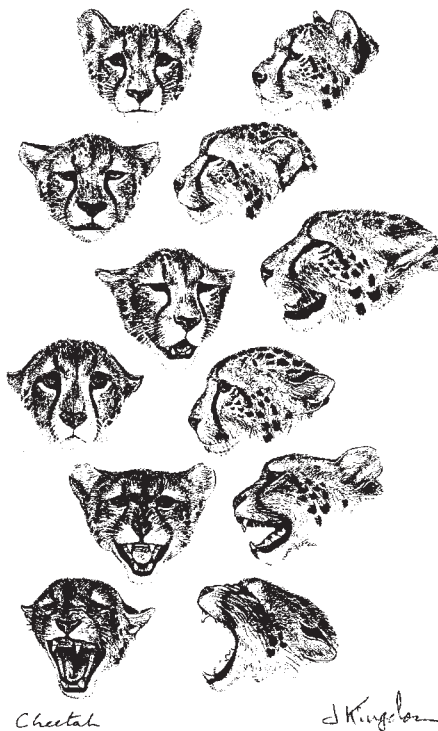
In 1983, O'Brien *et al.* created a stir with their paper announcing that cheetahs have remarkably little genetic variability^{1,2}. At that time, much research centred on questions of minimum viable population sizes, and particularly on how small a population could become before lack of genetic variability would make its long-term survival unlikely³. These days, considerations of genetic diversity are rarely at the top of any list of factors currently causing species endangerment^{4,5}, so social historians will not find it surprising that O'Brien's work has been subjected to two reappraisals^{5,6}. Both of the authors concerned, Caughley and Merola, argue that cheetahs are not especially impoverished genetically — and that it would be unlikely to matter much if they were.

The cheetah, *Acinonyx jubatus*, is found widely, if patchily, in sub-Saharan Africa, with a remnant population holding out in northern Iran⁷. The total cheetah population was estimated to be around 10,000–20,000 in the mid-1970s, having roughly halved over the preceding two decades as a result of hunting and changing land-management practices⁸. In Caughley's words, "how far they dropped further between 1975 and now is anyone's guess"⁵.

In their early work, O'Brien and colleagues found no heterozygosity at any of 47 isozyme loci in 55 cheetahs from southern Africa. Subsequent studies extended the analysis to two-dimensional electrophoresis of some 155 soluble proteins, finding a mean heterozygosity of $H = 0.014$ for East African cheetahs (H measures the fraction of loci that are heterozygous in the average individual). A previously undetected isozyme polymorphism was also found in southern African cheetahs, giving $H = 0.0004$ for this population. This figure contrasts with an average heterozygosity of $H = 0.071 \pm 0.002$ for vertebrates, or (more relevant) $H = 0.067 \pm 0.005$ for a collection of 172 mammalian species⁹. O'Brien and collaborators have concluded that the cheetah passed through a genetic bottleneck some 10,000 years ago, and that a second bottleneck is responsible for the exceptional loss of genetic variability in southern African cheetahs (possibly caused by the arrival of humans in this region only relatively recently)¹⁰.

Merola⁶ brings together evidence for levels of genetic variation for 24 other species of terrestrial carnivores, and finds eight that show no heterozygosity ($H = 0$). The coyote, *Canis latrans*, has an H -value above the southern African cheetah (H essentially zero), but below the East African ones. The remaining 15

species all have H -values higher than cheetahs, with the overall average for this set of 26 carnivores being 0.028. Merola thus argues that the genetic variability of cheetahs is more or less typical of carnivores. She further emphasizes that carnivores generally have less genetic variability than most mammals (she compares them with a group of 81 other



Variability of cheetah facial expression, depicted by Jonathan Kingdon. The genetic variability of cheetahs is a more contentious (but less visual) matter.

species, whose mean H value is 0.042; this contrast is even greater if we use $H = 0.067$ for a larger group of mammals⁹, including carnivores). Caughley⁵ independently makes broadly the same points, using much the same data.

In reply, O'Brien¹¹ emphasizes that Merola's and Caughley's comparison sets of H -values are based mainly on allozyme estimates of genetic diversity at a small number of loci. In particular, the assessments for the eight carnivore species rated by Merola as less genetically variable than cheetahs come from allozyme surveys of only 13 loci in the polar bear, *Ursus maritimus*, or 21 loci in the other seven species. The early cheetah studies of this kind used around 50 loci. But much electricity has flowed through gels since these early days, and O'Brien and co-workers' data are now based on more extensive studies, employing six additional measures: two-dimensional gel electro-

phoresis; acceptance or rejection of skin grafts among individuals (related to variability in the major histocompatibility complex, MHC); polymorphisms in length of restriction enzyme fragments within the MHC; similar studies of mitochondrial DNA; microsatellite polymorphisms; and fluctuating asymmetry in cranial measurements. Personally, I find the evidence from skin grafts and from fluctuating asymmetries a bit dodgy, but as a whole I think O'Brien's case is persuasive. The cheetah has markedly less genomic diversity than any other felid, including its closest relative, the puma (*Felis concolor*). The value of $H = 0.014$ is also low compared with other carnivore species studied in similar detail.

Even if the cheetah does indeed have an abnormally low degree of genetic variability, Merola and Caughley deny that there is much evidence of any deleterious effects in the form of inbreeding depression. O'Brien has argued that pronounced inbreeding depression shows up through high levels of abnormality in cheetah sperm, through litter sizes that seem comparatively small (averaging 1.5), through greater susceptibility to disease, and in general through the notable difficulty experienced by captive breeding programmes for cheetahs. Both Merola and Caughley contend that much of the difficulty in early captive breeding programmes was associated with differences between cheetah behaviour in the wild and in the constrained circumstances of zoos and preserves. Litter sizes in the wild range from 3 or 4 to as many as 5 or 6, and of 48 cub deaths reported in the wild, only one could have been attributed to genetic defects⁶. Nor are the sperm abnormalities off-scale for cats. Better management of programmes of captive breeding is leading to greater success.

Here I am inclined to side with the critics. Not least, there is other evidence that reduced genetic variability does not impair breeding success: today's healthy but relatively homozygous populations of Pere David's deer all spring from a very small founding population; and the bandicoot *Perameles gunnii* is widespread in Tasmania, despite its extreme homozygosity (interestingly, the relict population of fewer than 100 known individuals in mainland Australia shows quite high genetic variability¹²).

Both Merola and Caughley also argue that lack of genetic variability in the MHC may not put the cheetah at risk with respect to disease. O'Brien, however, points to an outbreak of feline infectious peritonitis virus in a wildlife park in Oregon; this mini-epizootic killed 19 cheetahs but none of the 10 similarly exposed lions. Merola dismisses this example on the grounds that the cheetahs were at much higher density than found in the wild. She goes further to speculate that infections of

this kind (microparasites, *sensu* May and Anderson¹³) may be less of a problem for carnivores, which, unlike most other mammals, tend to be solitary or occur as small groups at low densities; such diminished selection pressures could, she suggests, account for the systematically lower genetic diversity in the MHC in carnivores compared with other mammals. There is some sense in these speculations¹³. But they fail fully to appreciate O'Brien's main point, which is that relative homozygosity in the MHC — whether resulting from bottlenecks or, less likely, from selective considerations — can create problems with disease when such a population is brought into the relatively crowded conditions of captive breeding programmes, or when population densities increase in the wild. Cheetahs in Africa are indeed becoming more crowded: one cheetah per 6 km² in National Parks and other reserves, in comparison with a past figure of one per 100 km², according to one estimate⁸. So I think O'Brien's worries are well founded.

Merola's strongly expressed belief is that the cheetah's future is imperilled by loss of habitat and other consequences of human activities. She and Caughley see the lack of genomic variability as unimportant. O'Brien agrees — who could not? — that human effects are central to the cheetah's fate; he emphasizes, for instance, that cub mortality in the Serengeti is atypically high, partly because predator/prey ratios are higher there than in pristine settings, and partly because researchers working on cheetahs may inadvertently give other predators clues to the cubs' whereabouts¹¹.

In all, current thinking may rightly recognize that lack of genetic diversity is not the primary factor for most endangered species. But I share O'Brien's concern that it nevertheless remains an important consideration for many conservation programmes, and particularly for cheetahs. □

Robert M. May is in the Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK.

- O'Brien, S. J., Wildt, D. E., Goldman, D., Merril, C. R. & Bush, M. *Science* **221**, 459–462 (1983).
- O'Brien, S. J., Wildt, D. E. & Bush, M. *Sci. Am.* **254**, 84–92 (1986).
- Soule, M. E. (ed) *Viable Populations for Conservation* (Cambridge Univ. Press, 1987).
- Lande, R. *Science* **241**, 1455–1460 (1988).
- Caughley, G. J. *Anim. Ecol.* **63**, 215–244 (1994).
- Merola, M. *Conserv. Biol.* **8**, 961–971 (1994).
- Caro, T. M. in *Great Cats* (eds Seidensticker, J. & Lumpkin, S.) 138–147 (R. D. Press, Sydney, 1991).
- Myers, N. *IUCN Monographs* **4**, 1–90 (1975).
- Ward, R. D., Skibiuski, D. D. R. & Woodmark, M. *Evol. Biol.* **26**, 73–159 (1992).
- O'Brien, S. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 508–511 (1987).
- O'Brien, S. J. *Conserv. Biol.* **8**, 1153–1155 (1994).
- Robinson, N. A. *Conserv. Biol.* **9**, 114–125 (1995).
- May, R. M. & Anderson, R. M. *Nature* **280**, 455–461 (1979).

Bright spies for chiral molecules

A. Prasanna de Silva

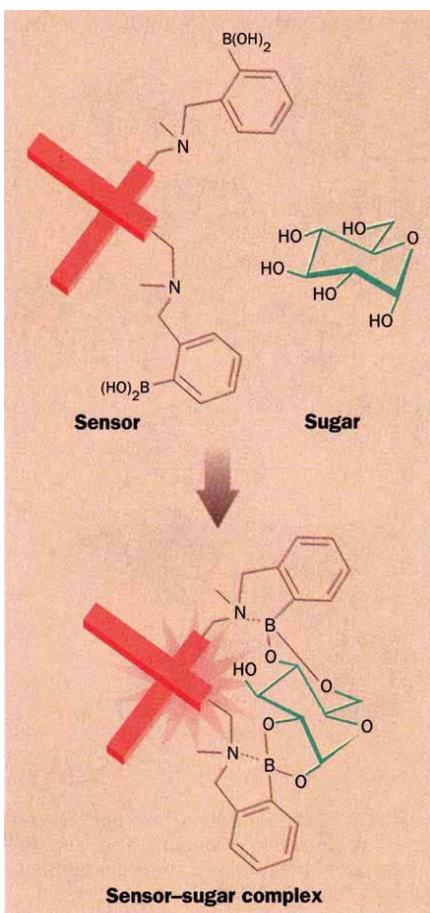
LIKE people, molecules can be left- or right-handed, and this chirality is especially common among biologically important molecules. For instance, D-glucose is essential for energy production in an organism whereas L-glucose (its enantiomer or mirror image) would pass un-

general approach, which includes more sophisticated techniques such as optical rotatory dispersion and circular dichroism, lacks the high degree of selectivity required for the monitoring of a given chiral molecule among many others in a complex matrix. This problem is especially serious for cases such as D-glucose because of its inability to absorb ultraviolet or visible light. One non-optical solution is to allow the natural enzyme glucose oxidase to act on D-glucose in the presence of air, and then detect the resulting hydrogen peroxide electrochemically². Because it is so important in the control of diabetes, the monitoring of D-glucose is now a profitable medical business.

A key to success in this and related methods is the recognition capability of the enzyme that selects D-glucose from the myriad of other molecules present in whole blood. In the work of James *et al.*, molecular recognition between sugars and phenyl boronic acids has been given a new twist with the insertion of a chiral, fluorescent 1,1'-binaphthyl as the backbone of the sensor¹. The inherent chirality of the sensor lets it distinguish one sugar enantiomer from the other. Selectivity for one sugar over other sugars is achieved by building two terminal boronic acids into the sensor for a pincer-grip.

The final part of the story is how the molecular event of selective recognition of D-glucose is translated into a humanly comprehensible message. Fluorescence is easily seen by the naked eye, can even be detected from a single molecule and is chemically switchable between 'on' and 'off' states³. Fluorescent sensors operating by photoinduced electron transfer (PET)⁴ exploit these features for the monitoring of many chemical species, including sugars^{5,6}. In many of these sensors, the fluorescence increases noticeably upon binding with the chosen analyte.

The present work of James *et al.*¹ belongs to this family but is rather special in that it introduces chirality into fluorescent PET sensing strategies for the first time.



A sensor of a given chirality binds preferentially with the sugar with the matching handedness and also produces a larger enhancement of fluorescence.

used through its digestive system. A more tragic example is thalidomide, where one enantiomer acts as a drug for morning sickness whereas the other leads to serious deformities in the fetus. A practical method of selectively monitoring chiral molecules of this type would therefore be most welcome. On page 345 of this issue, James, Sandanayake and Shinkai¹ demonstrate that they have made good progress towards this goal by building a pair of fluorescent sensors that discriminate between sugar enantiomers.

Optical monitoring of molecular chirality is not new. Indeed, a common manifestation of chirality is the rotation of the plane of polarized light by a solution of biologically active molecules. But this

- James, T. D., Sandanayake, K. R. A. S. & Shinkai, S. *Nature* **374**, 345–347 (1995).
- Schuhmann, W. & Schmidt, H.-L. *Adv. Biosensors* **2**, 79–130 (1992).
- Bissell, R. A. *et al. Chem. Soc. Rev.* **21**, 187–193 (1992).
- Bissell, R. A. *et al. Top. Curr. Chem.* **168**, 223–264 (1993).
- Yoon, J. & Czarnik, A. W. *J. Am. chem. Soc.* **114**, 5874–5875 (1992).
- James, T. D., Sandanayake, K. R. A. S. & Shinkai, S. *J. chem. Soc. chem. Commun.* 477–478 (1994); *Angew. Chem. int. Edn Engl.* **33**, 2207–2209 (1994).
- Irie, M., Yoroze, T. & Hayashi, K. *J. Am. chem. Soc.* **100**, 2236–2237 (1978).
- James, T. D., Harada, T. & Shinkai, S. *J. chem. Soc. chem. Commun.* 857–860 (1993).
- Tsien, R. Y. *Chem. Engng News* **72**(29), 34–44 (1994).

When two molecules of the appropriate handedness clasp together, the sensor's fluorescence is greatly enhanced (see figure). However, the beginnings of this development can be traced back to the enantioselective quenching of 1-1'-binaphthyl fluorescence first reported 17 years ago⁷.

In relevant earlier experiments, James, Harada and Shinkai combined sugar recognition by boronic acid with light interference in a twisted, cholesterol-based liquid crystalline phase⁸. This and related work hinged on a refreshing idea, but

sensory liquid crystal phases suffer from poorer resolution in space and time than truly molecular sensing strategies.

What of the future? We can expect enantioselective fluorescent sensing to be extended to make use of excitation and emission at longer, near-infrared wavelengths, eventually permitting monitoring of chiral molecules such as various L-amino acids and D-glucose in blood. Some of these sensors may find their way into diabetes monitoring in the home as electrochemical sensors have. Real-time fluorescence imaging of chiral molecules

within living cells may be another growth area in physiology, given the success of cyclic adenosine monophosphate (cAMP) imaging⁹. Also, regulatory pressure on the pharmaceutical industry to produce enantiomerically pure drugs should spawn new fluorescent sensors for testing the enantiomeric purity of various products and their metabolites. Overall, the prospects are bright. □

A. Prasanna de Silva is in the School of Chemistry, Queen's University, Belfast BT9 5AG, UK.

DEVELOPMENTAL BIOLOGY

Angles on activin's absence

Jim Smith

OVER the past five years, vertebrate embryologists have come to believe that activin, and the activin-binding protein follistatin, play a big part in the induction of mesoderm and neural tissue. Most of the evidence comes from *Xenopus*, where treatment of prospective ectodermal tissue with activin causes formation of mesodermal cell types¹, and inhibition of activin signalling — either by overexpression of truncated activin receptors² or by overexpression of follistatin³ — causes neural differentiation to occur. Activin and follistatin are indeed expressed during early *Xenopus* development: activin protein is present in the egg, perhaps provided to the oocyte by surrounding follicle cells^{4,5}, and follistatin is expressed in Spemann's organizer, the source of natural neural-inducing signals³.

Three papers that appear elsewhere in this issue⁶⁻⁸, beginning on page 354, now address the roles of activin and follistatin in mouse development. They show that although mouse embryos lacking activin, follistatin or a type II activin receptor are defective in several aspects of development, mesoderm formation and neural differentiation are not among them.

Activins are members of the transforming growth factor (TGF)- β family and function as dimers of two β -subunits, of which there are two homologous forms: β A and β B. Thus three types of activin have been isolated: activin A (β A β A), activin AB (β A β B) and activin B (β B β B). All three can induce mesoderm in the *Xenopus* assay⁹. In the mouse, both activin subunits are expressed zygotically before implantation¹⁰, and although embryonic expression declines after implantation, when mesoderm is formed, transcript levels are high in the deciduum (the maternal part of the placenta)¹¹.

The first investigation of the involvement of activin in mouse development was made by Jaenisch and his colleagues, who created mouse strains deficient in the β B

subunit¹². Homozygous mutant embryos had defects in eyelid development, and homozygous mutant females were unable to rear offspring normally. These were the only major defects, however, and the experiment showed unequivocally that mesoderm formation requires neither maternal nor zygotic expression of the β B subunit.

It remained possible that mesoderm formation in the mouse depends on activin A, and indeed the β A subunit is greatly upregulated in the ovaries of β B mutant females¹². To tackle this issue, Bradley generated mice mutant in the β A subunit and, with Jaenisch, also created animals that lack both subunits⁶. Mice lacking

the β A subunit die within 24 hours of birth. They have no whiskers or lower incisors, and they have a cleft palate, but mesoderm forms normally. Mice homozygous for mutations in both activin subunits also form mesoderm, and they show only the sum of the defects of the individual mutants. Mesoderm formation does not, therefore, require zygotic expression of any form of activin, and one activin subunit cannot compensate for the absence of the other.

Because mutant mice die within a day of birth, it was not possible to test whether females deficient in the β A subunit produce embryos lacking mesoderm. Those wishing to retain an interest in activin as a mesoderm-inducing factor might point out, therefore, that activin A derived from the deciduum might still act as an inducer. This is not as far-fetched as it might sound: maternal TGF- β 1 may cross the placenta to rescue embryos carrying a mutation in

Steps in time

FURTHER evidence for the diversity of the earliest land vertebrates comes from I. Stössel's report of this tetrapod trackway from the Upper Devonian of Ireland (*J. geol. Soc. Lond.* **152**, 407–413; 1995). The meandering track of a single animal, extending across a bedding plane of some seven by ten metres, was found in the Old Red Sandstone of Valentia Island on the southwest Irish coast. Although the individual prints are rather poorly preserved, the tracks as a whole suggests that the maker was a slow-moving, fully terrestrial amphibian with a stride of about a metre, and a length (between fore- and hindlimbs) of slightly more than this — the size of a small crocodile. Tetrapod tracks are extremely rare, and body fossils indicate that tetrapods were already highly diverse by the Famennian, the closing Stage of the Devonian Period, 360 million years ago. The scant fossil content of the Irish Old Red Sandstone makes age hard to assess, but a minimum age of Famennian for the prints seems likely. The fully terrestrial



nature of the maker contrasts with animals such as *Acanthostega gunnari*, the best-known early tetrapod, which is now thought to have been almost wholly aquatic. H.G.

this gene¹³, and it is equally possible that decidual activin A does the same thing. Moreover it is maternal, and not zygotic, activin that is required for mesoderm formation in the teleost fish *Oryzias latipes* (the Japanese medaka)¹⁴.

One way of asking whether maternal activin A might be responsible for mesoderm formation in the mouse is to interfere with the function of the activin receptor, as did Hemmati-Brivanlou and Melton in *Xenopus*². Activin receptors fall into two classes, both of which are serine/threonine kinases¹⁵. By analogy with the TGF- β receptor system, it is likely that the type II activin receptor is a constitutively active kinase. This molecule binds activin and then associates with and phosphorylates a type I receptor which goes on to initiate intracellular signalling¹⁶. If decidual activin A is required for mesoderm formation, a mouse embryo lacking a type II activin receptor would not be able to bind activin and should not form mesoderm at all. On the other hand, if decidual activin is not involved in mesoderm formation, such embryos should show a phenotype similar to activin-deficient mice.

It may not have been a surprise to Bradley and his colleagues⁷ to find that mouse embryos deficient in the type II receptor ActRcII form mesoderm perfectly well. But it is a surprise that defects in ActRcII-deficient mice show little resemblance to those of activin-deficient animals. Most individuals lacking ActRcII developed into adults whose only significant problem was the suppression of follicle-stimulating hormone (FSH) release. Activin stimulates FSH release¹⁷ (that is why it is called activin), and the knockout result suggests that this function is mediated solely through ActRcII. The embryonic functions of activin, however, must involve different or additional receptors, including, perhaps, ActRcIIB as well as receptors yet to be isolated⁷.

And what of follistatin? Do mice mutant for this gene lack a nervous system, as might be predicted from the work in *Xenopus*? The answer, to be found in the third paper⁸ in this issue, is 'no'. Follistatin-deficient mice are small, their muscle mass is reduced, their skin is shiny, and there are a number of other defects, but their nervous systems remain defiantly normal. Overall, the defects are more widespread than those seen in activin-deficient mice, implying, in agreement with its expression pattern¹¹, that follistatin may bind and modulate the functions of other members of the TGF- β family, or perhaps even function independently.

So this work⁶⁻⁸ counts against the view that activin and follistatin are involved in mesoderm formation and neural induction in the mouse, although activin addicts may still argue that decidual activin A

does the business, and those with a more general allegiance to the TGF- β family may point out that genes such as *nodal* are definitely required if mesoderm is to form properly^{18,19}. But what else do the new findings tell us?

First, they warn us that interpretation of experiments in *Xenopus* involving over-expression of gene products such as activin or follistatin or (especially) of truncated versions of activin receptors is fraught with difficulty. We need to know much more about what ligands and receptors are expressed at different stages of development and we need to know the biological specificities of the different molecules. As another warning, truncated type II activin receptors have been shown to block the action of the TGF- β family member Vg1 (ref. 20). In the case of follistatin, as well as this question of specificity, it is not even clear under what circumstances the protein binds to activin and inactivates it, and under what circumstances it 'presents' activin to a type II receptor.

But we should also be reminded that for all we learn from gene knockout experiments, what is lacking in the mouse is embryological data relating to mesoderm formation. In particular, we don't even know whether mesoderm in the mouse embryo really does arise through induction, still less when induction occurs and where the signal comes from. These embryological questions must be addressed at the same time as the analysis of mice mutant for putative 'mesoderm-inducing factors'. □

Jim Smith is in the Division of Developmental Biology, National Institute for Medical Research, The Ridgeway, London NW7 1AA, UK.

- Asashima, M. et al. *Wilhelm Roux Arch. dev. Biol.* **198**, 330-335 (1990).
- Hemmati-Brivanlou, A. & Melton, D. A. *Nature* **359**, 609-614 (1992).
- Hemmati-Brivanlou, A., Kelly, O. G. & Melton, D. A. *Cell* **77**, 283-295 (1994).
- Asashima, M. *Proc. natn. Acad. Sci. U.S.A.* **88**, 6511-6514 (1991).
- Dohrmann, C. E. et al. *Dev. Biol.* **157**, 474-483 (1993).
- Matzuk, M. M. et al. *Nature* **374**, 354-356 (1995).
- Matzuk, M. M., Kumar, T. R. & Bradley, A. *Nature* **374**, 356-360 (1995).
- Matzuk, M. M. et al. *Nature* **374**, 360-363 (1995).
- Nakamura, T. et al. *J. biol. Chem.* **267**, 16385-16389 (1992).
- Albano, R. M., Groome, N. & Smith, J. C. *Development* **117**, 711-723 (1993).
- Albano, R. M., Arkell, R., Beddington, R. S. P. & Smith, J. C. *Development* **120**, 803-813 (1994).
- Vassalli, A., Matzuk, M. M., Gardner, H. A. R., Lee, K.-F. & Jaenisch, R. *Genes Dev.* **8**, 414-427 (1994).
- Letterio, J. J. et al. *Science* **264**, 1936-1938 (1994).
- Wittbrodt, J. & Rosa, F. M. *Genes Dev.* **8**, 1448-1462 (1994).
- Kingsley, D. M. *Trends Genet.* **10**, 16-21 (1994).
- Wana, J. L., Attisano, L., Wieser, R., Ventura, F. & Massagué, J. *Nature* **370**, 341-347 (1994).
- Vale, W., Hsueh, A., Rivier, C. & Yu, J. in *Peptide Growth Factors and Their Receptors* (eds Sporn, M. B. & Roberts, A. B.) 211-248 (Springer, Berlin, 1990).
- Zhou, X., Sasaki, H., Lowe, L., Hogan, B. L. M. & Kuehn, M. R. *Nature* **361**, 543-547 (1993).
- Conlon, F. et al. *Development* **120**, 1919-1928 (1994).
- Schulte-Merker, S., Smith, J. C. & Dale, L. *EMBO J.* **13**, 3533-3541 (1994).

DAEDALUS

Crème de gaz

MANY oils and lipids form an emulsion with water. Some biological lipids, such as lecithin, go further: they form an emulsion of tiny hollow spheres full of water. These liposomes, as they are called, consist of one or more concentric bilayers of lipid molecules around an aqueous core. They can encapsulate drugs for delivery to specific sites in the body. Daedalus now points out that by standard surface-tension theory, any bubble is pressurized by its curved surface. The smaller its diameter, the sharper the curvature, and the higher is the resulting internal pressure. A liposome is a bubble only about 0.1 micrometres across, pressurized by all its curved surfaces in series. Its interior could easily reach 300 atmospheres.

So Daedalus is now loading liposomes, not with drugs, but with compressed gases. DREADCO chemists are pumping pressurized oxygen, acetylene, butane and so on into lecithin suspensions, stirring them vigorously to form liposomes, relaxing the pressure and churning the resulting liposomal 'milk' into a cream or butter containing 50 per cent or more of trapped gas.

'Gascream' will transform the bottled-gas industry. Out will go those clumsy cylinders. Instead, glassblowers, welders, anaesthetists and campers will simply spoon the appropriate Gascream into a special syringe and draw off the gas as it is released by thermal or chemical degradation of the cream.

But Gascream is more than a mere improved gas cylinder. An edible version pressurized with oxygen could make an ideal food for victims of pneumonia or bronchitis. Oxygen delivered via the stomach would save them from having to breathe. A perfectly balanced formulation could be metabolized completely in its own oxygen, making it the first fully self-contained food. Sadly, by the same token it would also be a chemical high explosive.

The DREADCO team is also creating unstable liposomes pressurized almost to bursting point: tiny disturbances will set them off. Liposomes loaded with high-pressure carbon dioxide could be the basis of a ferociously fizzy drink; a high-pressure liposomal shaving cream would give an unprecedented tingling-fresh sensation. Even pharmacy could benefit. Gas-laden liposomes could be a new weapon against bacteria or parasites. They would be the biological equivalent of the limpet mine. Once the organism had ingested or adsorbed the liposome, the slightest biochemical weakening would set it off. The invader would be blown apart in a microscopic blast of expanding gas. David Jones

Artefact or network evolution?

SIR — Enquist and Arak^{1,2} and Johnstone³ have chosen an inappropriate methodology to demonstrate principles of human perception of two-dimensional patterns in using neural nets with multi-dimensional input layers. When their input stimuli are displayed in two-dimensional arrays (*a, c* in the figure), the human eye sees symmetries and other spatial properties, but their artificial networks did not in fact use such spatial information in determining the correct response.

The input layers were actually *n*-dimensional (36-, 147- and 20-dimensional, respectively); their display as two-dimensional “retinal arrays” was therefore entirely arbitrary and had no relevance to network performance. The 6 × 6 input array in ref. 1 could equivalently be shown as a 4 × 9 array, and the 4 × 5 input array in ref. 3 could be displayed as a 5 × 4 array. If such alternative displays are chosen, the pattern labelled as a “conspecific bird” would appear to biological visual systems as remarkably un-bird-like (*b* in the figure) and the pattern labelled as the “most symmetrical tail” would appear remarkably asymmetrical (*d*). In other words, the networks did not learn anything about two-dimensional patterns — conspecific, symmetrical or otherwise — but they did learn the frequency with which individual input pixels indicated the correct output response. The simulation results in all three studies can therefore be explained entirely on the basis of the correlation coefficients between input and output units⁵.

In Johnstone's study³, the ϕ correlation coefficients⁶ between input and output units were strongly positive for all pixels in the central two columns in *c*. Whenever any such pixel was activated in trained networks, it was an indication that the pat-

tern was a “tail”, regardless of the two-dimensional configuration of the tail. For example, activation of pixel 2 indicates the presence of a “tail” in all five training patterns and never (or very infrequently, depending on the randomization technique) indicates a “non-tail”. This is why the most symmetrical tail (*c*) was recognized as a tail even when it was not included in the training set: all its pixels individually had strong correlations with the desired “tail” output.

Again, in Enquist and Arak's study¹, the pattern of ϕ coefficients rather than ideas about the perception of tail length explain the network results, whereas in their more recent simulation² point-biserial correlation coefficients⁶ between input and output units suffice to explain the evolution of symmetrical patterns. There, the nets were designed to select stimuli with strong primary colours, so that the rotation and reflection of input stimuli tended to produce strong synaptic weighting from input units located at positions 90, 180 and 270° from such inputs. The strong synaptic weighting in turn favoured selection of signals with the appropriate colour at corresponding signal positions, but these two-dimensional patterns, which biological visual systems see as symmetries, are, to such neural nets, nothing more than strong inputs in the 147-dimensional input space. Their positions in the input layer are irrelevant and their symmetry when displayed as a specific two-dimensional pattern plays no part in network performance⁵.

The problem in all three simulations is that neural networks are quick to exploit the correlational structure of the input-output vectors, whereas spatial forms rather than absolute retinal positions are the essence of perception by biological systems. As a consequence, although all the evolutionary arguments made in these three studies¹⁻³ may well be valid for other reasons, conclusions about the perception of two-dimensional patterns (objects, symmetries, tail-length, and so on) cannot be drawn on the basis of neural networks that have *n*-dimensional input layers.

Norman D. Cook

Faculty of Informatics, Kansai University,
Takatsuki, Osaka 569, Japan

JOHNSTONE REPLIES — Cook argues that the networks in my simulation³ did not learn anything about symmetry; rather, they merely learned the frequency with which individual pixels indicated the correct or incorrect output. Consequently, he suggests, the most symmetrical tail pattern elicited a strong response, not because of its symmetry (which played no role in network performance), but simply because it comprised pixels that were common to the set of tail patterns, and thus had a

strong positive correlation with the desired ‘tail’ output.

I entirely agree with this argument; in fact, I made the same point in my paper: trained networks preferred symmetrical tail patterns because these were close to the average of the training set, not because they were symmetrical *per se*, a point reinforced by the observation that random symmetrical images differing substantially from the training set average were not preferred.

The reason the symmetrical tail pattern was closest to the training set average (that is, comprised those pixels common to the largest proportion of training patterns) was that the asymmetric tail patterns in the training set formed mirror-symmetric pairs. Because paired ornaments in nature, such as swallows' tails or earwigs' forceps, also show symmetrical variation (FA being defined as deviations from symmetry that are random with respect to side), symmetrical males are also closer to the population average display. They, too, may be preferred simply for this reason, as a by-product of selection for mate recognition. This point is of some interest to evolutionary biologists, because it implies that preferences for symmetrical males can be explained without necessarily invoking a link between male symmetry and quality.

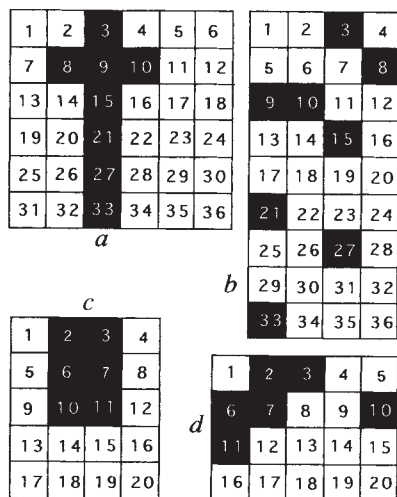
R. A. Johnstone

Department of Zoology,
University of Cambridge,
Downing Street, Cambridge CB2 3EJ, UK

ENQUIST AND ARAK REPLY — Cook suggests first that we have made the mistake of assuming that two-dimensional symmetry is information that is used by our nets²; and second, he raises a more general question of whether the particular neural nets we used really ‘see’ patterns in the same way that humans see them.

On the first point we agree with Cook that our neural nets did not code information about the geometrical structure of input patterns. This would be a serious criticism of our paper had we claimed that the networks could accomplish tasks of spatial perception, or detect symmetry *per se* in patterns, which clearly they cannot.

We started our simulations with an arbitrary pattern of input, and a pseudo-retina consisting of a number of cells arbitrarily arranged on a two-dimensional surface. As long as we assume the positions of input cells are fixed in space (as they are in a real retina), the conclusion that follows from our simulations is inevitable: in the process of selecting nets



a, “Conspecific male bird with long tail” when displayed in a 6 × 6 array; *b*, the same “bird” when displayed in a 4 × 9 array; *c*, the “most symmetrical tail” when displayed in a 4 × 5 array; *d*, the same “tail” displayed in a 5 × 4 array.

1. Enquist, M. & Arak, A. *Nature* **361**, 446–448 (1993).
2. Enquist, M. & Arak, A. *Nature* **372**, 169–172 (1994).
3. Johnstone, R.A. *Nature* **372**, 172–175 (1994).
4. Scalettar, R. & Zee, A. *Biol. Cybern.* **58**, 193–201 (1988).
5. Cook, N.D., Fröh, H. & Landis, T. *J. exp. Psychol.: Hum. Perc. Perf.* (in the press).
6. Kurtz, A.K. & Mayo, S.T. *Statistical Methods in Education and Psychology* 328–356 (Springer, New York, 1979).

for their ability to recognize random patterns in different orientations, networks evolve an increasing sensitivity to particular symmetries. Thus, by analogy, natural selection acting on organisms to recognize signals seen in different orientations offers a plausible source for some of the symmetries observed in biological signals (which evolve to become increasingly efficient at eliciting responses from the receiver). It is largely irrelevant whether we label our starting patterns as "bird", "tail" or otherwise¹; the point applies to any signal.

It is true that our results can be explained simply by correlations between inputs and outputs that develop when a pattern undergoes geometrical transformation on the pseudo-retina. But to regard our results as an "artefact" of such correlations distracts from the important question of whether such correlations would not also arise when a biologically significant pattern is projected in different orientations onto a real living retina. It seems difficult to avoid the conclusion that this will be the case.

Finally, we take issue with Cook's

implication that results obtained from artificial neural network studies are meaningful only if a net conceptualizes input patterns in the same way that humans (or other animals) do. To us at least, it seems questionable whether any model can do this. Although this may be a worthy aim for neurobiologists, who seek a description of the actual computational processes undertaken by the brain, as evolutionary biologists we use such models to explore how neural mechanisms may, in very general ways, influence the evolution of biological signals. In keeping with the best traditions of natural science, it would seem prudent to start with a simple model, rather than attempt to mimic the enormous complexity of real brains.

Anthony Arak

Archway Engineering,
Ainleys Industrial Estate,
Elland HX5 9JP, UK

Magnus Enquist

Department of Zoology,
University of Stockholm,
S-10691 Stockholm, Sweden

Knockout mouse fault lines

SIR — Despite the prevailing view that valuable information can be obtained with null-mutation experiments, concern has been expressed about the interpretation of the results in some biological systems (see refs 1–3). I believe that recent studies on memory now emerging from several different laboratories^{4–7} show that the null-mutant strategy, although a technological *tour de force*, is wholly inappropriate for resolving the issues for which it was intended. It is incapable of allowing the conclusion to be drawn that a particular molecule is necessary for an adult physiological process such as memory or long-term potentiation.

Put quite simply, null mutants are not animals without that protein, but are "reactionisms", organisms that respond to the mutation. This is so because the absence of a single gene is well known to alter expression of other genes and developmental programmes. These reactions to the 'knockouts'^{4–8} remain largely undocumented.

The mutation can be either with or without effect. In the case where the long-term potentiation or memory of the null mutant is altered, there is no compelling reason to ascribe the impairment to the loss of that protein when it is equally likely that the altered response results from the fact that the organism itself has been altered by the single gene mutation. In the case where the null mutation yields no loss of function, can one conclude that the molecule is not necessary for that function? I do not think so.

The phrase "considering the evidence for the apparently necessary and sufficient role of [protein x] in [function y] it is surprising that in the 'X' null mutant there is so little functional impairment" has become commonplace. Yet examples from the muscle differentiation field³ illustrate that a single knockout of an apparently critical transcription factor can be without effect. Because of documented increased expression in related genes in these null mutants leading to compensation, it would not be surprising if a similar event were also present in other single-gene knockouts. Therefore, one cannot conclude, when the protein but not the function is knocked out, that that protein is not necessary. Double knockouts now show that muscle differentiation no longer occurs³. Even here the conclusion that

Yet more danger for coelacanths

SIR — The living coelacanth *Latimeria chalumnae* is the sole survivor of the old lineage of crossopterygian fishes and is of great importance for evolutionary biology¹. Recently coelacanths were assigned to 'category 1' of the Convention of International Trade of Endangered Species (CITES). During the past three years we have witnessed an alarming population decline which we believe is due to fishing habits by local fishermen. We wish to alert the scientific community to the plight of the coelacanth and to suggest ways in which the apparent decline may be arrested.

We have been monitoring a coelacanth population, occupying 8 km of coastline off Grande Comore, for the past 6 years². Between 1991 and 1994 the average number of coelacanths in all underwater caves fell from 20.5 to 6.5 individuals. In 1991 we surveyed 59 individuals, but in 1994 only 40.

The fall in the number of individuals could be due to natural population fluctuation, emigration or even disturbances due to the presence of the submersible, but these possibilities seem unlikely. The decline is more probably attributable to human predation. Comorian fishermen, originally using paddle canoes, have regularly fished using long lines for food fish and for the oilfish (*Ruvettus pretiosus*) which live close to shore and at about the same depth as the coelacanth. Occasionally, they hook a coelacanth. Between 1989 and 1991 an almost stable number of coelacanths was sighted; this coincided with a fishing development programme. Fish-attracting devices were installed offshore, far outside the coelacanths' habitat, which dramatically reduced fishing pres-

sure on these animals. However, in 1994, our canoe survey revealed that fishermen, unable to afford repairs to their motorized canoes, were again forced to fish close to shore in the 'coelacanth zone'.

Because the CITES convention is signed by the Comorian government, it is strictly forbidden to land coelacanths, so not all accidentally caught coelacanths were reported. Further, our interviews revealed that several fishermen had killed coelacanths to retrieve their hooks.

We are concerned that the survival of the coelacanth, with an estimated population of about 200 individuals on Grande Comore², is severely threatened. We suggest that fish-attracting devices close to shore and within paddling distance, but set at well above the coelacanth zone, would be a useful fishing alternative. An information centre in one of the local villages and a permanently installed on-line underwater television system in front of one of the coelacanth caves could attract tourists and provide the community with income. The World Bank is interested in funding this project through the Global Environment Facility. International efforts should help to continue the population monitoring and scientific investigation of the coelacanth, whose survival depends not only on financial aid but on the education and cooperation of local communities.

Hans Fricke, Karen Hissmann

Jürgen Schauer, Raphael Plante

Max-Planck-Institut für

Verhaltensphysiologie,

82319 Seewiesen, Germany

1. Forey, P. L. *Nature* **336**, 727–737 (1988).

2. Fricke, H. & Hissmann, K. *Mar. Biol.* **120**, 171–180 (1994).

1. Neher, E. & Penner, R. *Nature* **372**, 316–317 (1994).

2. Malenka, R. C. *Nature* **372**, 218–219 (1994).

3. Weintraub, H. *Cell* **75**, 1241–1244 (1993).

4. Silva, A., Paylor, R., Wehner, J. & Tonegawa, S. *Science* **257**, 206–211 (1992).

5. Grant, S. G. N. *et al.* *Science* **258**, 1903–1910 (1992).

6. Abelson, A. *et al.* *Cell* **75**, 1253–1262 (1993).

7. Conquet, F. *et al.* *Nature* **372**, 237–243 (1994).

8. Strittmatter, S.M. *et al.* *Cell* **80**, 445–452 (1995).

9. Wagner, R. W. *Nature* **372**, 333–335 (1994).

these two transcription factors together are critical is open to question, because developmental distortions preceding muscle differentiation could have occurred.

If a function persists in the absence of a protein could one not conclude that, at the very least, the function can be performed without that protein? Yes, but given that many of the proteins being studied have isoforms or exist in a superfamily with shared motifs, the conclusion is meaningless if these isoforms take over that function. Null mutants are valuable models for study by developmental biologists interested in compensatory gene control mechanisms early in embryogenesis.

Because an enormous expenditure of time and research funds has been invested in the null-mutant strategy, it is important that constructive directions emerge from this technique. Perhaps the concerns raised here can be addressed by activatable promoters and region-specific knock-outs. Implementation of other approaches such as those exploiting anti-sense techniques⁹, which avoid developmental issues but are not without their own problems, may be of potential use. Whatever the strategy selected, it must take into account that the organism does not passively receive the experimenter's manipulation.

I think that the question of the necessity of a gene product cannot be answered with the null-mutant preparation. Because of the accelerating rate in which these mutations are being carried out I am worried that the study of the molecular basis for information storage is being led into a technological cul-de-sac, as the data generated are essentially uninterpretable.

Aryeh Routtenberg

Departments of Psychology and
Neurobiology and
Institute for Neuroscience,
Northwestern University,
Evanston, Illinois 60208, USA

Distinguishing truth from lies

SIR — In the report of the 'Megalab Truth Test' (*Nature* 373, 391; 1995), R. Wiseman concedes that "it was not possible randomly to allocate individuals to the three conditions" of receiving information via television, radio or newspaper. Indeed, anyone having a slight acquaintance with British society and media would expect the populations tapped by BBC1 television, Radio One, and *The Daily Telegraph* to differ systematically in their distribution of age, education, social class, degree of cultural commonality with Sir Robin Day and, quite possibly, sex. Those responding to the question must have been self-selected within each audience by further criteria (which led, for instance, to

a much smaller percentage response from the radio audience than from the other two groups). Given that no attempt was apparently made, either in the experimental design or in the analysis, to control for these variations, it would be rash indeed to draw any conclusions about the role of verbal, vocal and visual cues in lie detection.

The Megalab was intended to increase public understanding of science. In the kind of social and behavioural science exemplified by this study, the public need little stimulus to understand the interest of the questions. However, the broader public, students in these subjects, and even some of *Nature's* readership in the other sciences, need a great deal of education about the problems, pitfalls and possibilities in reaching scientific conclusions from statistical information. It is a pity that the unique opportunity provided by Megalab was used to offer an example of a social psychological experiment in which these problems were not seriously confronted.

Oliver Braddick

Department of Psychology,
University College London,
Gower Street, London WC1E 6BT, UK

Wing shape in pterosaurs

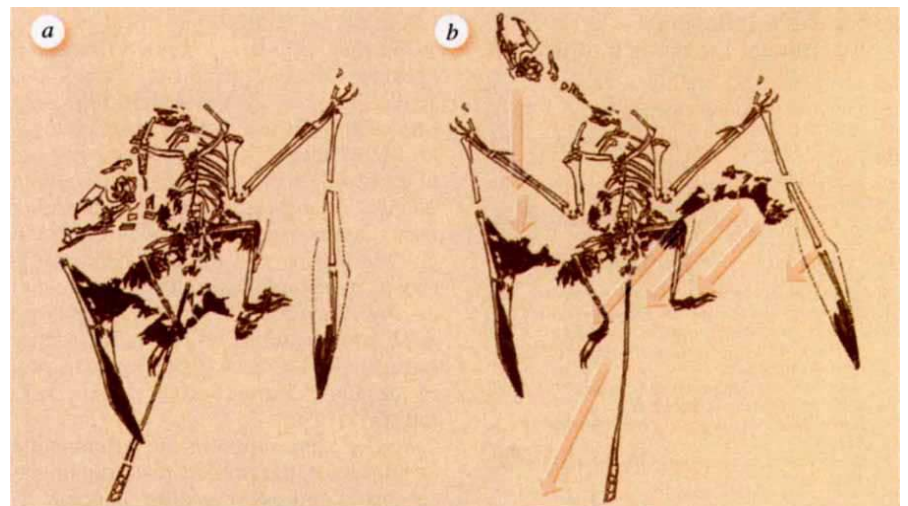
SIR — No consensus has yet been reached on the shape of the wing membrane in pterosaurs. A specimen (PINM 2585/3) of the middle Asian pterosaur *Sordes*^{1,2} seems to present compelling evidence for a broad bat-like wing membrane attached laterally as far as the ankle, and for a medial membrane spanning the hindlimbs and lateral digits. Reconstructions of the membranes using pen and ink¹ or half-tone screens² imply homogeneity and continuity up to the hindlimbs, but in my view

the published photographs do not support those implications.

Large comet-tail smears of tissue near the ankles and feet (*a* in the figure) are clues that parts of the decaying carcass have drifted from their original positions. The tail was 'weathervaned' in the same direction as the comet-tails, and the left forelimb and head had also moved in the same direction. Although not immediately apparent, the left forelimb, partly hidden beneath the skull, had also drifted back in the same direction. The specimen thus exhibits more post-mortem disturbance than the best Solnhofen specimens (refs 3, 4; in contrast to ref. 2), but it apparently preserved hairs better. There is no evidence of bottom currents in the prehistoric lake in which the remains were found^{5,6}, but the fall of the specimen from the lake's surface and its impact on the bottom could account for these disturbances.

I have created a reconstruction (*b*, from a published colour photograph⁷), pushing elements back to their original symmetrical positions. Missing parts of the left wing were mirror-imaged from the right side. Despite the replacement of the missing wing phalanx, symmetry was preserved only when the trailing edge of the left wing shifted forward to the knee area. In doing so it no longer terminated at the ankle as reported^{1,2}. The smears were moved, along with the median membrane, to their apparent points of origin. I placed these on the wing because the thighs were apparently well preserved and complete.

Further evidence that the medium membrane was originally part of the wing comes from the observation that the 63° angle present in the trailing edge of the "uropatagium" almost matches the 65° angle present in the trailing edge of the left wing. In both wing tips, strong parallel fibres preserve a non-extensible feather-like shape. The middle, extensible portion of the left wing is apparently ragged and falling apart. Perhaps in the right wing the



a, Tracing of *Sordes pilosus* (PINM 2585/3) (from ref. 7). *b*, The same specimen partly reconstructed. Arrows indicate probable direction of drifting elements.

process was advanced to the point that it had become detached. In the photographs (Fig. 1a of ref. 2), neither wing membrane clearly indicates where it was attached to the leg, although illustrations^{1,2} of the right wing show an imprint identical in shape and position to that of the left wing. This observation may be suspect because, as reconstructed here, the left wing may have originally attached near the knee.

I can accept that a membrane may have spanned from the lateral digits to either side of the tail (it would not impede terrestrial locomotion), but I question how a membrane spanning the legs and attached medially can also be attached to the lateral digits. The image of pterosaurs with broad bat-like wing membranes is traditional, but at least two Solnhofen specimens^{3,4,7} were preserved with very narrow wings stretched only between the elbow and wing finger. Thus, either *Sordes* is different from these specimens, or previous interpretations^{1,2} are in error.

David Peters

12812 Wood Valley Court, St Louis,
Missouri 63131-2051, USA

UNWIN AND BAKHURINA REPLY — Reconstructions of the Upper Jurassic pterosaur *Sordes pilosus* with broad wings attaching to the hindlimbs and a medial membrane, the uropatagium, between the legs^{1,2}, have been challenged by Peters, and others⁴, on the grounds that preservation of the remains is too poor to permit such inferences. However, comparison of *Sordes* with other exceptionally preserved pterosaurs shows that the holotype is one of the best-preserved individuals in existence^{2,8}. As originally interpreted, the skeleton was complete and fully articulated, with the neck flexed backward as in some Solnhofen pterosaurs⁹. Each forelimb is partially folded and, although parts of the left forelimb were lost during collection, it is evident that, like the hindlimbs, they are almost perfectly symmetrical with respect to each other, contradicting claims⁴ of post-mortem disturbance.

Interestingly, the position of the skele-

ton corresponds almost exactly to that exhibited by complete, undisturbed examples of *Rhamphorhynchus*, from the Upper Jurassic Solnhofen limestone¹⁰. However, the evidence for soft tissues, especially the extent and structure of the wing membranes, is much clearer in *Sordes*. The wing membranes of Solnhofen pterosaurs are preserved as impressions^{3,4,9,10}, but even the best examples can be interpreted in a variety of ways^{3,4,11}. The decomposition of *Sordes* was halted at an earlier stage than in Solnhofen pterosaurs: extensive tracts of black, mineralized soft tissues⁸ pick out impressions left by the wing membranes. Outlines of the wings can be traced with ease and show, quite unambiguously, the existence of a uropatagium and attachment of the cheiropatagium to the leg as far as the ankle^{1,2,8}.

Peters' narrow-winged reconstruction of *Sordes* is based on a highly unreliable technique, interpretation of photographs, and unfounded suppositions of post-mortem disturbance. Central to his argument is the claim that the uropatagium is part of the right cheiropatagium which, somehow, drifted into a symmetrical position between the hindlimbs, the internal fibres fortuitously lining up parallel to the lower leg. However, Peters' interpretation (his figure b) omits parts of the central region of the right cheiropatagium, preserved lateral to the right knee and clearly visible together with the uropatagium in

Sharov's¹ plate 4. Impressions of the flight membrane, the posterior boundary of which is evident in Sharov's¹ plate 5, Fig. 1a, show that the right cheiropatagium remained intact and did not undergo any post-mortem transport. The wings occupy almost identical positions, as Sharov correctly showed¹, but the outline of the right cheiropatagium is more difficult to trace in photographs because of the patchy preservation of the black, mineralized tissues in this area.

Peters' reconstruction also fails to explain the unusual morphology and position of the fifth toe^{2,8,12} and is contradicted by details of the internal composition of the membranes. The uropatagium, preserved in three other individuals as well as the holotype, contains relatively short, sinuous, loosely packed fibres, clearly distinguishing it from the middle and outer parts of the cheiropatagium, which is characterized by much longer, straighter, closely packed fibres².

David M. Unwin

Department of Geology,
University of Bristol,
Bristol BS8 1RJ, UK

Natasha N. Bakhurina*

Palaeontological Institute,
Russian Academy of Sciences,
Moscow 117647, Russia

*Present address: Department of Geology, University of Bristol.

Early AGEing and Alzheimer's

SIR — Mattson *et al.* suggest¹ that in Alzheimer's disease "glycation is a late event" and that it "results from free radical generation". Despite their citation of some of our papers to support their view²⁻⁴, their opinion is directly contrary to the findings that the protein PHF- τ /A β 8, regarded as the precursor to neurofibrillary tangles⁵, is modified by advanced glycation end products (AGE)^{2,6}.

We have also demonstrated that AGE epitopes localize to diffuse amyloid- β senile plaques³ (lesions that represent one of the earliest pathological changes in Alzheimer's disease⁷), as well as neurofibrillary tangles and neuritic senile plaques. Moreover, the proteins found in these lesions are continuously exposed to glucose *in vivo* and therefore constantly subject to Schiff base formation followed by Amadori rearrangements and AGE modification. Of relevance is that intracellular AGE modification, contrary to popular dogma, often occurs very rapidly⁸.

Rather than representing a tombstone or late event, the process resulting in glycation-related modification appears to represent a very early event that promotes the formation of free radicals⁹ and

enhances the aggregation of τ or amyloid- β ⁹ into neurofibrillary tangles and senile plaques. Therefore, a contribution of glycoxidation modifications early, as well as later, in the pathogenesis of Alzheimer's disease should be seriously considered.

Mark A. Smith*

Lawrence M. Sayre*

Michael P. Vitek†

Vincent M. Monnier*

George Perry*

*Institute of Pathology,
Case Western Reserve University,
Cleveland, Ohio 44106, USA

†The Picower Institute,
Manhasset,
New York 11030, USA

- Sharov, A.G. *Akad. nauk SSSR paleont. Inst. Tr.* **130**, 104–113 (1971).
- Unwin, D. M. & Bakhurina, N. N. *Nature* **371**, 62–64 (1994).
- Wellnhofer, P. *Ann. Naturhist. Mus. Wien* **88A**, 149–162 (1987).
- Padian, K. & Rayner, J. M. V. *Am. J. Sci.* **293A**, 91–166 (1993).
- Gekker, R. F. *Trudy paleont. Inst.* **15**, 7–85 (1948) (in Russian).
- Doludenko, M. P., Sakulina, G.V. & Ponomarenko, A.G. *The Geology and Late Jurassic Fauna and Flora of a Unique Locality, Aulie (Karatau Range, Southern Kazakhstan)* (Inst. Geol., Russian Acad. Sci., Moscow, 1990) (in Russian).
- Wellnhofer, P. *The Illustrated Encyclopedia of Pterosaurs 1–192* (Salamander, London, 1991).
- Bakhurina, N. N. & Unwin, D. M. *Hist. Biol.* (in the press).
- Wellnhofer, P. *Abh. Bayer. Akad. Wiss. (N.F.)* **141**, 1–133 (1970).
- Wellnhofer, P. *Palaeontographica* **149**, 1–30 (1975).
- Pennycuik, C. J. *Biol. Rev.* **63**, 209–231 (1988).
- Bakhurina, N. N. & Unwin, D. M. *J. vert. Paleont.* **12** (Suppl.), 18A (1992).

- Mattson, M. P., Carney, J. W. & Butterfield, D. A. *Nature* **373**, 481 (1995).
- Yan, S. D. *et al. Proc. natn. Acad. Sci. U.S.A.* **91**, 7787–7791 (1994).
- Smith, M. A. *et al. Proc. natn. Acad. Sci. U.S.A.* **91**, 5710–5714 (1994).
- Vitek, M. P. *et al. Proc. natn. Acad. Sci. U.S.A.* **91**, 4766–4770 (1994).
- Trojanowski, J. Q. & Lee, V. M. *Am. J. Path.* **144**, 449–453 (1994).
- Ledesma, M. D. *et al. J. biol. Chem.* **269**, 21614–21619 (1994).
- Giaccone, G. *et al. Neurosci. Lett.* **97**, 232–238 (1989).
- Giardino, I., Edelstein, D. & Brownlee, M. J. *clin. Invest.* **94**, 110–117 (1994).
- Hunt, J. V. & Wolff, S. P. *Free Radic. Res. Commun.* **12–13**, 115–123 (1991).

Capturing the light

David H. DeVorkin

The Perfect Machine: Building the Palomar Telescope. By Ronald Florence. Harper-Collins: 1994. Pp. 434. \$27.50.

Stairway to the Stars: The Story of the World's Largest Observatory. By Barry Parker. Plenum: 1994. Pp. 332. \$27.95.

In 1928, George Ellery Hale mused that "Starlight is falling on every square mile of the earth's surface, and the best we can do at present is to gather up and concentrate the rays that strike an area 100 inches in diameter". Much of the recent history of astronomy has been about the quest for more light-gathering power, and these two books, written for general audiences, reflect this drive.

Starting in the late eighteenth century with Herschel's speculum reflectors, whose mirrors ranged up to four feet in diameter, and accentuated in the 1840s with Lord Rosse's six-foot 'Leviathan of Parsonstown' mirror, astronomical telescopes have tended to become larger, more sophisticated and very expensive. Save for the efforts of a few astronomers such as Herschel and Rosse, however, large refractors dominated the professional landscape until the first part of the twentieth century, when speculum reflectors gave way to silver-on-glass mirrors and the eye as detector gave way to the photographic plate. Both changes were due not only to new technologies but also to new motives among professional astronomers. No longer content to measure the positions of the stars or the appearances and motions of the planetary bodies in the Solar System, astronomers wanted to probe the depths of space in order to unravel the mysteries of the nebulae and the compositions of celestial bodies. They therefore started using spectroscopic techniques that were insatiable consumers of light.

Although not alone in the field — though Ronald Florence would have the reader believe otherwise — Hale still takes the credit as the consummate entrepreneur for establishing the large reflector as the instrument of choice in astrophysical research. He was, actually, representative of a transitional era, when the grand long-focus refractors were still being built but were being accompanied by reflectors at the same site. Observatories at Paris, Toulouse and Marseilles, as well as those at Greenwich and Cambridge in England, at the Lick Observatory in California and at Hale's own Yerkes Observatory in Wisconsin, which centred on a 40-inch refractor, all boasted silver-on-glass reflectors up to four feet in diameter by 1900. But then Hale took the next step, opening a 60-inch in 1908 at his new mountain site in southern

California, the Mount Wilson Solar Observatory. The 60-inch stood on its own, without a great refractor, and was not built for solar work. In fact Hale never confined his solar observatory to work on the Sun, but saw solar studies as part of a triad embracing laboratory and stellar spectroscopy integrated to answer the great problem of the sidereal universe: the origin and evolution of the stars.



Astronomical greats — Edwin Hubble in the prime focus cage of the 200-inch telescope at Palomar.

Hale, however, was not driven to build larger telescopes to solve specific problems. He was in too much of a hurry for that. The 60-inch was barely operational before he and his staff were planning the next step, the 100-inch Hooker reflector, which opened on the same site in 1917 and was fully operational the next year. And only a few years after that, while the Hooker was yielding the modern picture of the structure of the Universe — that galaxies exist and are moving away from one another — Hale and company dreamed of another leviathan, in the 200–300-inch range. With the opening of the 200-inch reflector on top of Mount Palomar in 1948, a decade after Hale's death, his direct influence ends, but not his legacy.

Indeed, one gets the strong impression from *Stairway to the Stars* that Hale's ambitions are alive and well on the barren Moon-like landscape of Mauna Kea on the island of Hawaii. For this is home to a growing population of high-technology wonders of astronomical engineering: one admires both the industry and ingenuity of the modern entrepreneurs as they build one telescope after another, each larger or more specialized than the last, with the latest giant the dual 10-metre Keck I and II segmented mirror telescopes.

Parker does a pedestrian but credible job of describing the basic characteristics of this modern phalanx of light collectors operated by groups from Hawaii, mainland universities, NASA (National Aeronautics and Space Administration) and institutions from Canada, France and the United Kingdom. His narrative account, based on visits to the site and interviews with astronomers, physicists and engineers, plainly reveals how the infrastructure of modern astronomy has become highly complex, international and, to a great extent, political.

Most astronomers who build telescopes, he reveals through formulaic capsule biographies, are really physicists looking for something interesting to do. They are the ones who have come up with many of the technological and social shifts that distinguish the modern giant telescope from the now venerable 200-inch: with only a few exceptions, no individual, government or philanthropic agent today will foot the bill for a new leviathan. Consortia of universities and countries typically are called upon, or created, to do the job.

The observatory funded by the Keck Foundation is the partial exception to which Parker pays most attention, although curiously he does not look too closely at the foundation's philanthropic motives or the many fistfights during the complex technical history of the telescope. He does however demonstrate that the astronomical talent of two great institutions in California — the California Institute of Technology and the University of California — were required to design, build and operate the telescope system, which is still under development. When he wrote his book, Keck I was just coming on line, and its twin 10-metre, Keck II, had just been approved. Overall one gets the impression that Mauna Kea is the best place in the world to do astronomy. Accordingly, the five rave reviews on the dust jacket all happen to be by astronomers from Hawaii.

Parker's book is not a history of telescope building: Hale's name is never men-

tioned, nor is his great achievement noted except in passing. Florence makes a grand attempt to bring both back in his flamboyant but flawed study of the building of the 200-inch reflecting telescope named for Hale at its dedication. The book is flamboyant because he takes great pains to provide gobs and gobs of colour commentary. He paints a grand landscape of the building of the telescope, starting with criticisms of Hale's ambitious dreams in the late 1920s and the way in which the Rockefeller Foundation was won over, to disasters in casting a suitable mirror blank, first at General Electric and then at Corning in the 1930s, to the myriad engineering problems that had to be overcome in constructing the telescope's mounting. Here is the most intriguing part of the story. The experimental styles of Caltech and Westinghouse engineers differed over how to build such large structures as the huge horseshoe-shaped polar bearing. They were never fully reconciled, nor were Westinghouse's billings completely forthright. And even when the telescope was completed, Ira Bowen's staff had many difficulties getting the telescope to work right in the late 1940s. Although he gets the overall story straight, Florence fills in this landscape with incredibly luscious details historians would die for if they were from credible, retrievable sources: revealing motives and inner thoughts, perceptive conversations, provocative impressions of the action, political and personal assaults, with all the intrigue of a Faustian drama. But therein lies the flaw.

A writer of both fact and fiction, Florence knows how to spin a yarn, but unfortunately he confuses the two in the details that fill the early parts of *The Perfect Machine*. He pulls in bits and pieces from a wide range of sources, sometimes citing them, more often not. But he fails in places to be critical of his sources, such as reminiscences, and gets himself into trouble when he over-uses them. For example, the first two chapters outline in florid detail the famous session at the National Academy of Sciences when Harlow Shapley and Heber Curtis came to an impasse over "the scale" of the Universe. Although he cites it elsewhere in his book, Florence does not reveal that he depended very heavily on Shapley's famous but inaccurate *Through Rugged Ways to the Stars* for his description, in Chapter 2, of the proceedings of the session. Shapley stated correctly that the session was in April 1920, but he conflates Einstein's well-known 1921 visit with his 1920 session, and tells a colourful but imaginary story accentuated by Einstein's presence, with samplings from Einstein's wit at dinner that night. Florence uncritically repeats Shapley's tale, including Einstein and his wit at the scene. To make it all fit, Florence then puts the

Shapley-Curtis encounter in 1921, repeating the error in his table of contents, chapter title and text.

This blunder, and its probable cause, demonstrates the author's sympathy for sensation over critical analysis, which makes it difficult to trust the rest of the book, no matter what the enthusiastic endorsements say on the book's jacket. For instance, Florence then goes on to portray the impasse between Shapley and Curtis as a stimulant for Hale's thinking about larger telescopes. Without citation, he expresses Hale's inner thoughts during the National Academy session about how the impasse convinced him that what the world really needed was a new huge telescope to solve the questions raised at the meeting. There is no more proof for this than there is for Einstein's presence at this event.

Ironically, the doubt created by what

appears to be an unnecessary invention lingers throughout what is, intrinsically, a very interesting and important story. It is ironic because Florence later recounts the charges of invention and inaccuracy levelled by astronomers at David O. Woodbury's *The Glass Giant of Palomar*, published in 1939, calling it "a lively though sometimes fanciful book". Florence, sadly, has failed to learn from Woodbury's example. So although the book reads well and is engaging, especially the later sections, one must read it with care, almost as if it were a historical novel, which is a disservice to that venerable and still active telescope, to its builders and to the real legacy of Hale. □

David H. DeVorkin is at the National Air and Space Museum, Smithsonian Institution, Washington DC 20560, USA.

How to be an astronomer

David Hughes

The Guide to Amateur Astronomy, Second Edition. By Jack Newton and Philip Teece. Cambridge University Press: 1995. Pp. 335. £24.95, \$34.95.

The Observer's Guide to Astronomy, Volumes 1 and 2. Edited by Patrick Martinez. Cambridge University Press: 1994. Pp. 1,148. Vol. 1: £50, \$79.95 (hbk); £24.95, \$34.95 (pbk); Vol. 2: £50, \$79.95 (hbk); £24.95, \$34.95 (pbk).

Compendium of Practical Astronomy, Volume 1: Instrumentation and Reduction Techniques; Volume 2: Earth and Solar System; Vol. 3: Stars and Stellar Systems. Edited by Günter Dietmar Roth. Springer: 1994. Vol. 1: Pp. 540, DM98 (pbk); Vol. 2: Pp. 362, DM68 (pbk); Vol. 3: Pp. 321, DM68 (pbk).

ASTRONOMY is one of the most accessible subjects for the amateur. Go outside on a clear night, look up into the sky and you will have begun. Even suburban dwellers will see a few hundred stars, and soon the questions will start. Why do the stars have different brightnesses and colours? Why are some close together whereas others are in large loose groups? Who could have possibly joined up the stellar dots into such convoluted constellation patterns and then conjured up such a marvellous array of animals and gods to represent these constellations?

Soon the star-gazer becomes dissatisfied with the naked eye and starts to clamour to see more and more stars and ever-increasing detail on planetary and satellite surfaces. This introduces the first serious problem. Should he or she invest in a pair of binoculars or in a telescope? Should the telescope have an objective lens or a reflecting mirror? How big does

the telescope need to be and should it be placed somewhere convenient or somewhere dark? And what should be looked at and how useful will the observations be to other people?

Fortunately, Jack Newton and Philip Teece's book contains the answers to these frequently asked questions. *The Guide to Amateur Astronomy* leads the new observer step by step from the easy stellar and planetary objects to the more difficult sight-seeing. The book bubbles with enthusiasm; the authors have clearly enjoyed every hour they have spent at the telescope and their astronomical excitement is infectious. The reader is introduced to such topics as the joys of solar eclipses and comet searching, the challenge of sketching planetary surface details and estimating the varying brightnesses of eclipsing stars and the skills of nebulae hunting and astrophotography. The authors even go so far as to discuss the building of simple observatories, the making of one's own telescope, essential darkroom techniques and the choice of software for one's home computer.

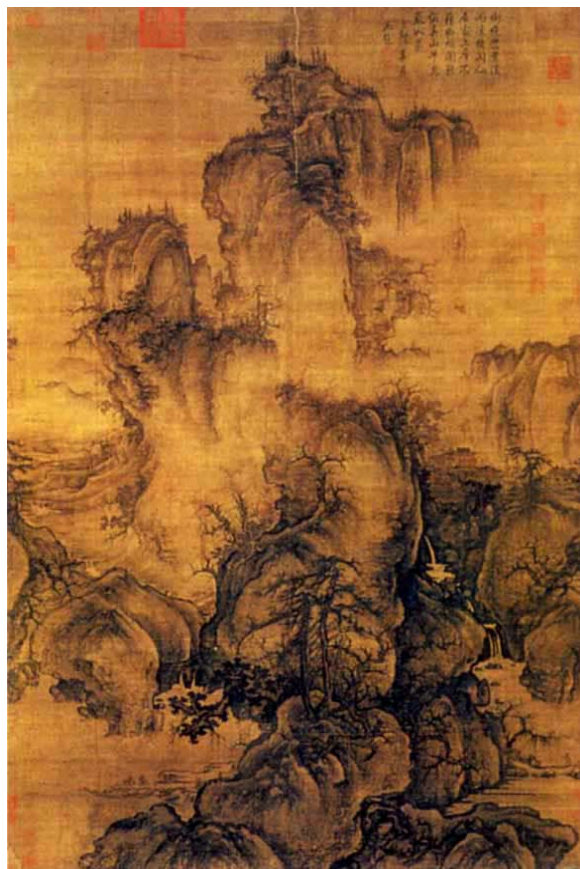
The Observer's Guide to Astronomy is written for the advanced amateur astronomer. There is more to astronomy than just revelling in the beauty of the celestial scene; the advanced observer must sooner or later get on with doing something useful. Luckily, amateur astronomy and professional astronomy are truly complementary. Although the professional astronomer may have access to large instruments and fast computers, amateurs can cover fields that need huge amounts of observing time, inexpensive equipment, mobility and an impressive number of observations. For example, annual monitoring of the activity of the

Perseid meteor shower needs hundreds of skilled amateur astronomers all over the world. Their combined results will not be over-troubled by limited hours of darkness or bad weather at any one site.

In *The Observer's Guide to Astronomy*, a work first published by the Société d'Astronomie Populaire in 1987, Patrick Martinez edits a collection of 20 contributions by experienced French amateur astronomers. Not only are the practical methods of observation dealt with in great detail but the scientific background is also stressed. The aim is to train astronomers to make better and better observations, and to guide them towards the correct analysis of these observations and to the most useful way of presenting the data. The emphasis is on the development of good observational and analytical skills.

Whereas the first two titles are for the amateur and the advanced amateur astronomer, *Compendium of Practical Astronomy* goes one step further. We are now in the realms of the accomplished amateur, the university observatory, the teacher and the budding professional. The book's predecessor, *Handbuch für Sternfreunde*, was first published in 1960, and the English edition was published in 1975. We now have a superbly translated and slightly expanded version of the fourth edition in which 21 professional astronomers contribute 28 chapters covering the length and breadth of practical astronomy. This is an invaluable source-book, detailed and thorough: it contains, for example, a two-page table on the physical characteristics of the materials used to construct telescopes and their mountings, as well as circuit diagrams for frequency-control units for synchronous telescope motor drives and ten criteria for successful telescope mounting. Also discussed are astrophotography, the fundamentals of spectral analysis, the principles of photometry, error analysis, sundial construction and the correct approach to the history of astronomy. A host of observational projects are suggested for such diverse objects as the Sun, the Moon, artificial satellites, comets, asteroids, meteors, planetary surfaces, noctilucent clouds, aurorae, variable stars, binary stars, galactic features and extragalactic objects. Great care has been taken to inform observers about how best to present and analyse their data. The book also contains a host of useful tables and figures that help the reader to compare data from different sources. The illustrations are first class and there are extensive references. I have nothing but praise for this book. It was an astronomical classic more than 30 years ago and is now even better. □

David Hughes is in the Department of Physics, University of Sheffield, Hicks Building, Sheffield S3 7RH, UK.



"Early Spring" (1072) by Kuo Hsi. In his "Essay on Landscape Painting", this Chinese writer, philosopher and painter discusses four types of landscape art: those in which one could travel, gaze upon, dwell within or ramble. For him nature is alive: "Water-courses are the arteries of a mountain; grass and trees its hair; mist and haze its complexion". The picture is reproduced from the sumptuously illustrated *A Vision of Nature: Traces of the Original World*, in which Michael Tobias provides a highly personal — and sometimes irritatingly mystical — look at the "aesthetic, psychological and philosophical impact of Earth on humanity". Kent State University Press, \$39.

In Nature's infinite book

David E. Allen

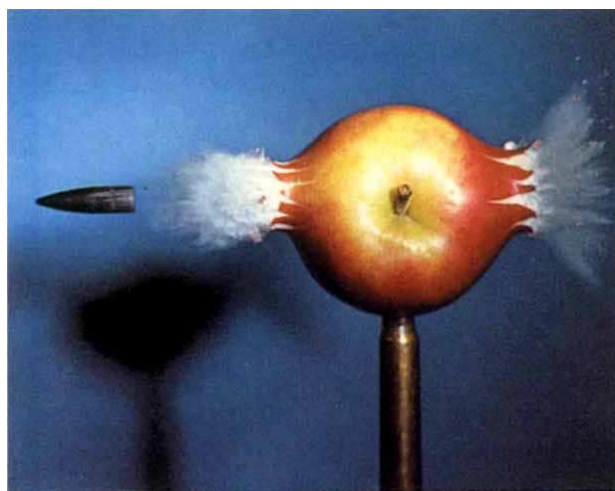
The Oxford Book of Nature Writing. Edited by Richard Mabey. Oxford University Press: 1995. Pp. 249. £16.99.

It was almost as predictable as that grass will grow that Richard Mabey, Britain's foremost present-day writer on the English countryside, would eventually succumb to an anthology. Surprisingly, although half a century ago there was one devoted just to birds and another, more recently, to the pursuit of butterflies and moths, natural history as a whole has largely passed untouched by the genre. This may be because works of this type were long regarded as too chancy commercially; more certainly, they demand a wide knowledge of the literature combined with a confident personal taste, qualifications not all that common in a field that tends to throw up either narrow specialists or the belletristically inclined insufficiently ballasted with factual information. Just as in nature writing itself, there is a necessary middle way that all too many find elusive.

Sensibly, Mabey has kept his choice

within manageable bounds by confining himself to the Western tradition and just to writings in prose, excluding both poetry and fiction. This has enabled him to concentrate on how naturalists themselves — for the most part — have described what they have seen and recorded. Even so, it has not prevented him from giving rein to one or two indulgences, such as the phrenological calendar of an eighteenth-century Swede, a bare list of one plant's English vernacular names and a classification of the animal world by a tenth-century Chinese author — this last a double bit of cheating. More ambitiously, he has further selected his material with the aim of illustrating how attitudes to nature have altered over the centuries, grouping it under several sections designed to reflect various broad tendencies of a more or less vaguely historical kind, within each of which the individual pieces are arranged roughly chronologically. Prefacing each section is a brief historical scene-setting by himself, which adds much to the appeal of the whole.

Most authors embarking on such a work would have been tempted to give undue coverage to zoology and to birds and mammals in particular (the concern, after all, of far and away the greatest number of today's students of natural history). To his credit, however, Mabey has insisted on including a proper ration of botany,



Stopping time — Harold "Doc" Edgerton's stop-action stroboscopic flash photographs of the path of a bullet through an apple and the spiked coronet of a milk-drop are just two of the many famous images that poured out of MIT's "Strobe Alley", the hallway where he worked and taught. He is shown here with his friend Jacques Cousteau (right), the underwater explorer. An account of Edgerton's life and science is given in *Seeing the Unseen* edited by Roger R. Bruce, which, together with a photo CD, presents more than 250 images of the work of this inspiring inventor, teacher and entrepreneur. George Eastman House, \$53.95, £35.95 (distributed by MIT Press).

revealing in the process the very wide range that constitutes one of his special strengths. In his choice of extracts, too, he reveals himself as laudably catholic: few others, surely, would have thought of, or even known of, John Aubrey's jottings on English trees, for instance. At times, though, he goes too far: Cowper's observations on the behaviour of his cats belongs more rightly in an anthology of petkeeping.

Starting with Aesop and Aristotle and ending with a piece written just a decade ago by Primo Levi, there is something here for everyone, from Gilbert White (predictably) and Charles Darwin and Thoreau to such less likely people as Ruskin, Colette and Cyril Connolly. Botanists will be pleased to find gems by the too-little-known Andrew Young and Jocelyn Brooke, not just examples of the evergreen Anne Pratt and Flora Thompson. Similarly, for bird-watchers there are Edmund Selous and J. A. Baker to add some out-of-the-way spice to the familiar W. H. Hudson and Richard Jefferies. Natural history in the United Kingdom receives perhaps disproportionate attention, but the topics are not entirely neglected and US readers will be reassured to some extent by the inclusion of

John Muir and George Perkins Marsh (though not, slightly surprisingly, Ernest Thompson Seton). The range of possibilities is enormous, but Mabey has passed over impressively few of the favourite authors that more immediately come to mind. Edward Newman is perhaps the most gaping omission.

If one were to criticize the selection at all, it might be on the ground that there is too little on the social side of fieldwork, on how the natural-history community actually ticks. Very rightly and properly, we are given an account of a fungus foray in 1869 from the *Transactions* of the Woolhope Club of Hereford, England, but I would have welcomed some excerpts from its near-contemporary, the *Entomologists' Intelligencer*, or from Loudon's *Magazine of Natural History*. In the pages of journals such as these one will not find jewelled prose, but they bubble with an ingenuous enthusiasm of a kind that is rarely otherwise captured on paper but which has always been of the subject's essence. There is more to natural history than nature itself. □

David E. Allen is in the Wellcome Institute for the History of Medicine, 183 Euston Road, London NW1 2BE, UK.

Top prize for science writer

The British-born cosmologist and popular science writer Paul Davies has won the £650,000 Templeton Prize for Progress in Religion, the world's largest annual monetary award. Davies, aged 48, is professor of natural philosophy at the University of Adelaide and has held academic appointments at the universities of London, Cambridge and Newcastle upon Tyne. He emigrated to Australia in 1990. As author of more than 20 books, among them *God and the New Physics*, *The Mind of God* and *The Last Three Minutes*, he has achieved an international reputation as a gifted popularizer of science.

Previous winners of the prize have included Alexander Solzhenitsyn, Mother Teresa, the evangelist Billy Graham, the Watergate burglar-turned-preacher Charles Colson and the scientists Alister Hardy, Carl Friedrich von Weizsäcker and Charles Birch. Founded in 1972 by the financier Sir John Templeton and administered in New York, the award is given to a living person who has shown "extraordinary originality" in advancing humankind's understanding of God and spirituality.

Davies says that "science cannot prove the existence of a design or designer [but] it can reveal the sheer depth of ingenuity that goes up to make this marvellous universe, our home. Scientific research will continue to illuminate theological issues, and no religion which ignores these advances can hope to remain credible".

He left the United Kingdom partly because of what he called "the Thatcher government's deliberate and systematic butchery of science". Ironically, one of the judges of this year's prize was Lady Thatcher. Davies intends to use the award to pay for his own new chair.

New in paperback

The Refrigerator and the Universe: Understanding the Laws of Energy by Martin Goldstein and Inge F. Goldstein. Harvard University Press, \$15.95, £12.75. See *Nature* **368**, 598 (1994).

The Language Instinct by Steven Pinker. Penguin, £8.99. See *Nature* **368**, 360 (1994).

The Second Law: Energy, Chaos, and Form by P. W. Atkins. Revised edition. Scientific American Library/W. H. Freeman, £14.95.

Weather Cycles: Real or Imaginary? by William James Burroughs. Cambridge University Press, £14.95, \$19.95.

Guide to the Sun by Kenneth J. H. Phillips. Cambridge University Press, £13.95, \$15.95.

Animal Behaviour, Second Edition by Mark Ridley. Blackwell Science, £16.50.

The consequences of imperfect mixing in autocatalytic chemical and biological systems

Irving R. Epstein

When chemical reactions whose rate increases with the concentration of a product species are carried out in imperfectly mixed systems, a variety of complex behaviours can occur. These phenomena, which have relevance for biological processes as well, include chaotic and stochastic behaviour and selection of one final state over an equally probable alternative.

ONE of the great conceptual revolutions that has taken place in the physical and, to a lesser extent, the biological sciences in this century is the recognition that nonlinearity is ubiquitous and that it has profound effects on the dynamics of a system¹. Although linear systems, in which superposition principles and the intuitively pleasing notion that the whole is equal to the sum of its parts, are often susceptible to exact mathematical analysis, they constitute but a small fraction of the real world. More and more, particularly with recent advances in computational power, science deals with nonlinear systems that give rise to complex temporal oscillation, chaos and spatial pattern formation.

One point that has perhaps been underappreciated is the extent to which spatial inhomogeneities can interact with and amplify the temporal nonlinearities in an evolving system so as to yield new and important phenomena. In any system governed by a nonlinear rate law, a knowledge of the bulk or average concentrations is not sufficient to predict the average rate of the reaction. One must also know the spatial distribution of reacting material. The further the rate laws depart from linearity, the more drastically the behaviour of the system changes when the spatial distribution is modified.

In many chemical and biological systems undergoing reaction and diffusion in the absence of mixing, strikingly similar and rather beautiful patterns spontaneously develop. The time-dependent spiral and target patterns seen in the Belousov-Zhabotinsky reaction², in aggregating slime moulds³ and in developing oocytes⁴ have been the subject of much attention, as have the stationary patterns of spots or stripes predicted by Turing⁵ as a mechanism for morphogenesis and recently found experimentally in simple chemical systems⁶.

Here I focus on the dramatic effects of a different aspect of the interaction between temporal nonlinearity and spatial inhomogeneity, the behaviour of imperfectly mixed, autocatalytic systems. Imperfect mixing in autocatalytic systems can, for example, lead to chemical reactions that occur at seemingly random intervals, crystallization processes that yield sometimes all left-handed and sometimes all right-handed crystals, and reactions in which changing the rate at which a solution is stirred can cause a transition from a stationary, time-independent state to one of periodic or even chaotic oscillation. These phenomena have been studied thus far primarily in inorganic chemical media; but there are possible implications also for living systems.

Autocatalysis

Autocatalysis—the enhancement of the (absolute) rate of growth of a population as that population increases—is ubiquitous in biological systems. It is, in a sense, the essence of reproduction. As the Reverend Thomas Malthus recognized when he suggested that population tends to grow geometrically, the number of individuals born is, on average, proportional to the number in the population. This perceptive clergyman was also aware that there

are limitations to population growth; he proposed⁷ that the food supply grows only arithmetically. As an alternative, one might construct a model of population growth that includes a fixed carrying capacity K for the environment, so that the population $N(t)$ tends toward K as time $t \rightarrow \infty$

$$dN/dt = \beta N(1 - N/K) \quad (1)$$

where β represents the average *per capita* excess of the birth rate over the death rate. Equation (1) is the logistic equation, one of the simplest and most thoroughly studied models of population dynamics. It contains the crucial elements of interest here, autocatalysis (the first term on the right-hand side) and nonlinearity (the second term, proportional to N^2).

Autocatalysis occurs also in non-living systems, albeit less commonly. In chemical systems it can lead to explosion (analogous to “the population explosion”) if, for example, an exothermic reaction with a rate that increases sharply with temperature is run in a closed vessel that does not permit the heat generated by the reaction to escape rapidly enough.

Autocatalytic chemical reactions, when coupled to other reactions that provide a negative feedback to control the proliferation of the autocatalytic species, can give rise to such complex behaviour as periodic oscillation or chemical chaos. In configurations open to the flow of matter and/or energy, such as stirred-tank reactors⁸, autocatalytic chemical systems can support multiple stable states, leading to hysteresis as a control parameter is varied.

Bifurcations and mixing

In many studies of chemically reacting media the assumption is made that the system is “well mixed”, that is, that the concentration of each species is uniform throughout the system. To bring about this presumed state of homogeneity, experimenters often employ mechanical agitation, for example, a magnetic stirrer that revolves at hundreds or thousands of revolutions per minute (r.p.m.). If the reactions under study are nonlinear, any major departure from the assumption of perfect mixing can have major dynamical consequences.

In fact, perfect mixing is almost never achieved in practice, and chemical engineers, who must deal with the consequences of imperfectly mixed reactors containing large quantities of nonlinearly reacting materials, have been aware of this situation for decades. Mixing is an extremely complicated phenomenon occurring on multiple length scales and timescales. Engineers speak of two aspects of mixing, macromixing and micromixing. Macromixing is the intermingling of packets of fluid, such as those that enter a flow reactor in different input streams, to achieve a composition that appears homogeneous if sampled on a macroscopic scale. Micromixing, which is far slower and more difficult to achieve, involves the segregation of a fluid in vortices and filaments at very short length scales. Although much pro-

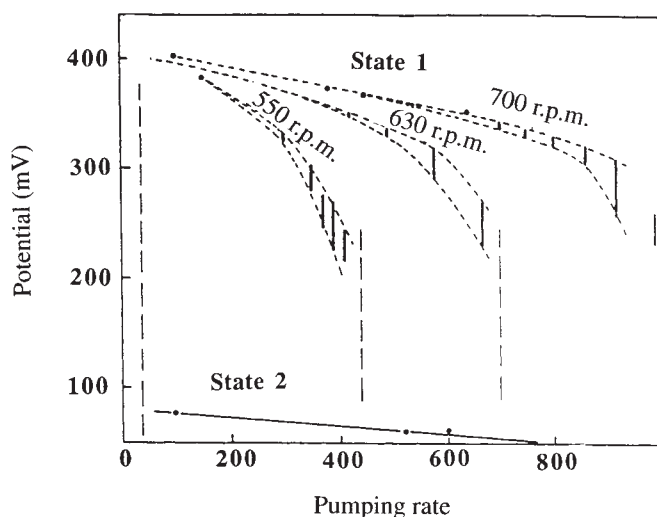


FIG. 1 Effects of stirring rate on bistability in the chlorite-iodide reaction¹⁰. Dashed line, steady-state potential in state I; solid line, steady-state potential in state II. Note how the bistable region shifts and broadens as the stirring rate is increased from 550 to 700 r.p.m. (The pumping rate is given in units of $\sim 0.02 \text{ ml min}^{-1}$; the potential is that of a Pt electrode versus an $\text{Hg/HgSO}_4/\text{K}_2\text{SO}_4$ reference electrode.)

gress has been made toward understanding the factors that govern the degree of mixing in a system and toward developing reactor designs that improve the quality of mixing, particularly macromixing⁹, the goal of a truly well-mixed system remains a distant one.

Rather than pursue the quest for the elusive (perhaps unattainable) grail of perfect mixing, it may be more useful to recognize the extent of imperfection that exists in systems of interest, and to assess the consequences that such imperfection brings.

One of the first demonstrations of the importance of mixing quality is found in a set of experiments by Roux *et al.*¹⁰ on the chlorite-iodide reaction, a prototype system in nonlinear chemical dynamics¹¹, in a stirred-tank reactor. Typical results are shown in Fig. 1. Although the bistability and hysteresis in this system had previously been characterized both experimentally and theoretically, it came as a shock to most investigators that the critical flow rate for transition from the upper to the lower branch of steady states was strongly dependent on the stirring rate, even at relatively rapid rates. Luo and Epstein¹² later demonstrated that the stirring rate can be treated as a bifurcation parameter, like the flow rate or the input concentration of a reactant, which shifts the dynamical behaviour of the system among a series of steady and oscillatory states.

A simple model¹³, illustrated in Fig. 2, provides a qualitative understanding of the origin of the phenomenon. Inevitably, some regions of the reactor, for example those closest to the stirrer, are better mixed than others. As the stirring rate increases, the relative size of the poorly mixed volume shrinks, although it never reaches zero. As a first approximation, we represent the actual reactor by two coupled reactors: one, the well-mixed region with volume V_a , in contact with the input and output flows; the other, the poorly mixed region with volume V_d , linked only to the well-mixed region but not to the external flows. The parameters Q_x (the cross-flow between V_a and V_d) and the ratio V_a/V_d are used to characterize the quality of the mixing, which improves as the stirring rate increases.

The originators of the model suggest that it is the ratio of the transition time between steady states of differing composition to the residence time in the reactor that determines how changing the degree of mixing will influence the stability of the two bistable branches of steady states. If the transition time (induction period) is relatively short, then poorly mixed regions will

be found on the thermodynamic (low-flow-rate) branch, that is, with concentrations similar to those of the completed reaction. Decreasing the stirring rate will tend to stabilize this state at the expense of the high-flow-rate branch. If, in contrast, the induction period is relatively long, isolated regions will tend to be found whose concentrations resemble those of the unreacted input flow—material will be “washed out” of the reactor before it has a chance to react to any significant degree. Decreasing the stirring rate under these conditions will tend to decrease the range of stability of the low-flow-rate branch.

This crude model, which may be thought of as the first in a hierarchy of models that characterize the macromixing process in increasing detail, gives surprisingly good agreement with the experimental results, not only in the chlorite-iodide reaction, but also in systems whose chemistry causes the mixing quality to affect primarily the high-flow-rate rather than the low-flow-rate branch. The branch that is most sensitive to mixing quality and the sequence of bifurcations that occurs as the stirring rate is varied can be accurately predicted, but a quantitative connection between the model parameters on the one hand and the stirring rate and reactor geometry on the other remains elusive. In a real reactor, of course, one does not have two homogeneous zones of different composition, but rather a composition that varies continuously throughout the reactor volume.

A more quantitative picture of the degree of departure from perfect mixing at various stirring rates in the chlorite-iodide reaction was provided by Menzinger and Dutt¹⁴. They placed a platinum microelectrode at different distances from the central stirring rod in their cylindrical reactor. In one experiment, at a stirring rate of 880 r.p.m., consistent with many experiments on such systems, they observed a difference in electrode potential of about 50 mV when the electrode was moved from a position at the axis to a position 12 mm away in the radial direction. This difference corresponds to a factor of roughly eight in the

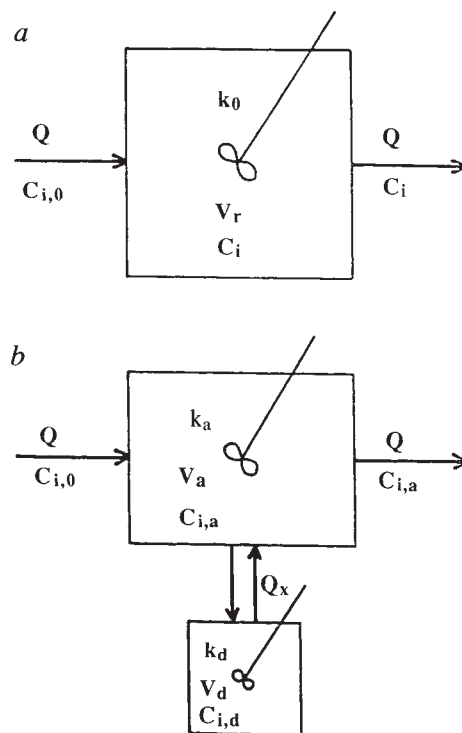
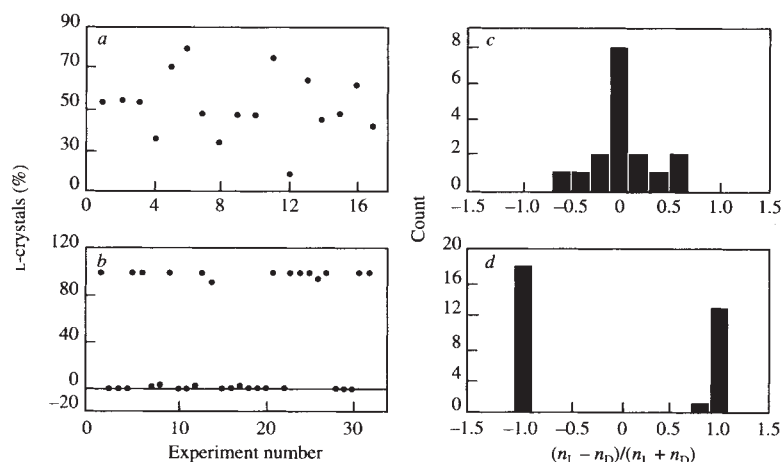


FIG. 2 Two-reactor model¹³ of imperfect mixing in a stirred-tank reactor. *a*, Idealized system in which mixing is perfect and concentrations are homogeneous throughout the reactor. *b*, Stirred tank with cross flow between active (well mixed) and dead (poorly mixed) zones. Symbols used: Q , volume flow; C , concentration; k , flow rate (reciprocal residence time); V , volume. Subscripts: *i*, chemical species; *o*, input from reservoirs; *r*, homogeneous reactor; *a* and *d*, active and dead zones, respectively.

FIG. 3 Statistical analysis of crystallization from stirred and unstirred sodium chlorate solutions²⁶. a, b, Scatter plots of the percentage of L-crystals for unstirred (a) and stirred (b) experiments. c, d, Enantiomeric excess of L- over D-crystals in unstirred (c) and stirred (d) experiments.



iodide concentration! Even at a stirring rate of 2,820 r.p.m., far above the rates normally employed, and very near the threshold for cavitation, $[I^-]$ differed by 17% between the axis of the reactor and the tip of the stirrer.

Some effects of imperfect mixing

Chaos and mixing. The two-reactor model shown in Fig. 2 played a major role in a recent controversy about the origins of chemical chaos in the Belousov–Zhabotinsky (BZ) reaction. Although by the mid-1980s chaotic behaviour had been found experimentally by several groups under a variety of conditions^{15, 17}, no chemically realistic model had been shown to generate chaos. In 1989, Györgyi and Field¹⁸ demonstrated that if the oregonator model¹⁹ for the BZ chemistry was incorporated into a two-reactor model, chaos was easily obtained. They then argued that the experimentally observed chaos arose, not from the inherent (homogeneous) chemical kinetics of the BZ system, but from the coupling of those nonlinear rate laws with imperfect mixing in the experiments.

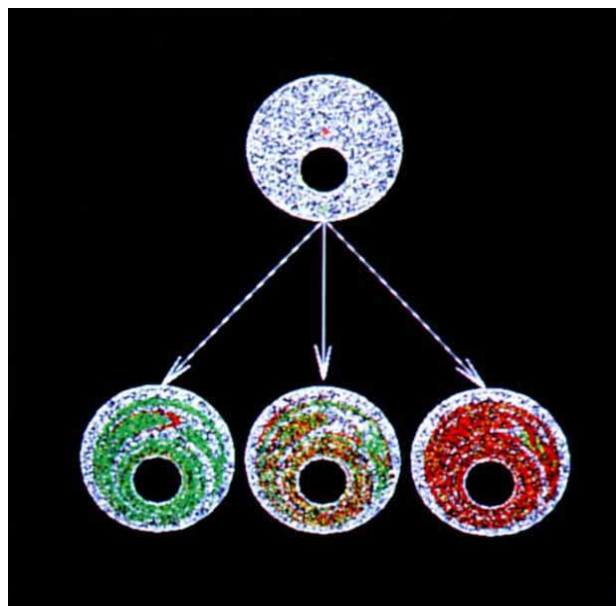
The Györgyi–Field proposal called into question the source of chaotic behaviour in the best characterized chaotic chemical

system. After several months of occasionally heated discussion with experimentalists, Györgyi and Field developed a chemically plausible variant of the oregonator model that produces chaotic behaviour with perfect mixing under conditions resembling those in the experiments^{20, 21}. Although it remains extremely difficult to pinpoint the origin of chaos in any particular experiment, it now appears that chemical chaos in the BZ reaction can result either from the homogeneous chemistry or from perturbation of that chemistry under periodic oscillating conditions by the effects of imperfect mixing.

Experimental evidence of chaos has since been found in other chemical reactions, such as chlorite–thiosulphate²², chlorite–thiourea²³ and oxidase–peroxidase²⁴. Even where model calculations²⁵ imply that homogeneous chaos is possible, these systems are as yet too poorly characterized for us to be able to state unambiguously whether or not mixing plays an essential role in the observed chaos.

Chiral symmetry-breaking. Questions about origins, whether of the Universe, of the Earth or of life, have always held great fascination. One intriguing problem of this type is to explain chiral symmetry-breaking in nature; how, for example, it came

FIG. 4 Simulations of particles in a continuum of solute reacting according to equation (2). The green seed starts at the same location in all three simulations. The initial position of the red seed is moved by 3% to give the three different final states shown in the bottom row. Each simulation starts with 7,000 unreacted white particles, one red and one green seed. With different initial conditions, the flow can produce any final ratio of red to green particles. (By courtesy of G. Metcalfe.)



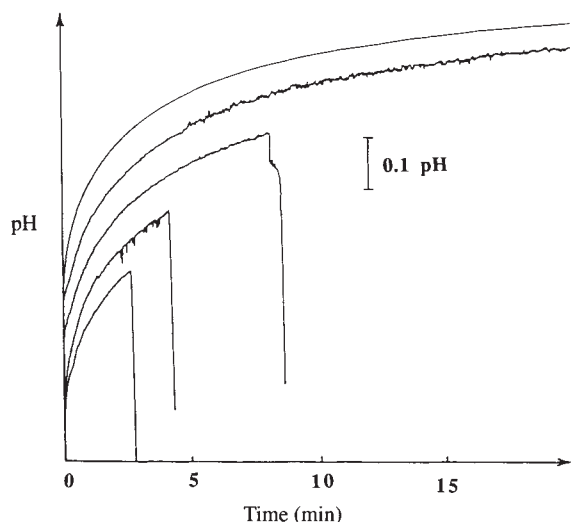


FIG. 5 Representative pH traces in the chlorite-thiosulphate reaction³⁰ at 20.0 °C measured under identical conditions. Successive curves have been shifted by 0.07 pH units for easier viewing, because in the absence of a shift, the initial portions of the curves coincide nearly perfectly.

to be that nearly all the optically active amino acids of biological significance are laevo- rather than dextro-rotatory. In an experiment that rivals Pasteur's work on optical activity for its simplicity and elegance, Kondepudi *et al.*²⁶ obtained some remarkable data on the distribution of optical activity among crystals precipitating from a supersaturated solution of sodium chlorate under different mixing conditions. The handedness of the crystals was established visually, using crossed polarizers.

The results are summarized in Fig. 3. When the solution was unstirred, of 1,000 crystals collected from 17 different crystallizations, 525 were of laevo (L) and 475 of dextro (D) chirality; that is, the total numbers were statistically equal, as were the numbers in each individual experiment, as seen in Fig. 3a and c. When the solution was stirred at 100 r.p.m. during the crystallization process, the total numbers for the 11,829 crystals were consistent with a 50-50 distribution of D- and L-crystals. With only one exception, however, each of the 32 individual crystallizations (Fig. 3b and d) gave rise to a set of crystals that was more than 99% laevo- or 99% dextro-rotatory!

The authors attribute the observed selectivity to an autocatalytic effect arising from the rapid production of secondary nuclei from a single primary nucleus in the stirred system. In effect, the first "product" particle, the primary nucleus or microcrystal, serves as a chiral template on which further products, the secondary nuclei that grow into crystals, are formed. More recently, Kondepudi and Sabanayagam²⁷ have demonstrated by scanning electron microscopy that one mechanism of secondary nucleation in sodium chlorate involves the formation and subsequent breaking off by stirring-induced shear forces of needle-shaped structures on primary crystals of sufficient size. They suggest an expression for the rate of secondary nucleation that takes into account the need for the supersaturation, the size of the primary nucleus and the stirring rate all to exceed their threshold values.

Metcalf and Ottino²⁸ have modelled the experiments of Kondepudi *et al.*²⁶ by a simple autocatalytic reaction scheme combined with a chaotic mixing flow. The 'chemistry' is mimicked by a pair of autocatalytic reactions involving white (W), red (R) and green (G) particles:



The mixing is approximated by a bounded, two-dimensional, well characterized, Stokes flow characteristic of fluid residing between two parallel but non-concentric cylinders²⁹. The simula-

tion is started with a large number of white particles and a single autocatalytic seed particle of each colour. Although the simulations do not lead to 'ultra-high-purity product' of the sort found in the experiments on sodium chlorate crystallization, many initial conditions quickly yield states in which over 90% of the coloured particles are of the same colour (Fig. 4). Perhaps more suggestive of the extreme sensitivity of this system is the fact that very small changes in the initial location of one of the coloured seeds can turn a red-dominated state into a green-dominated one. The authors point to the detailed topology of the mixing, the "highly interwoven nature" of the manifold along which the dominant colour propagates, as the source of this surprising behaviour.

Crazy clocks. Another phenomenon in which autocatalysis and imperfect mixing interact with dramatic effect involves reactions in which there is a sharp transition from an initial, largely unreacted state to the final, equilibrium state. A number of autocatalytic reactions that occur in aqueous solution and that give rise to a colour change serve as the basis for lecture demonstrations known as 'clock reactions'. In a clock reaction, reactant solutions of known concentration are mixed, and there is apparently no reaction until, suddenly, perhaps accompanied by a wave of the hand or magic words from the demonstrator, the solution changes, for example, from red to blue. The time at which the colour change takes place is predictable to very good accuracy (the 'clock' aspect) from a knowledge of the reactant concentrations and the temperature.

In some clock reactions, however, there is a relatively narrow range of concentrations in which the time of this sharp transition becomes essentially unpredictable. The prototype system of this type is the chlorite-thiosulphate reaction³⁰. Measurements of pH versus time for five replicate experiments starting from the same initial concentrations are shown in Fig. 5. For the first several minutes, the curves are nearly identical. The pH increases

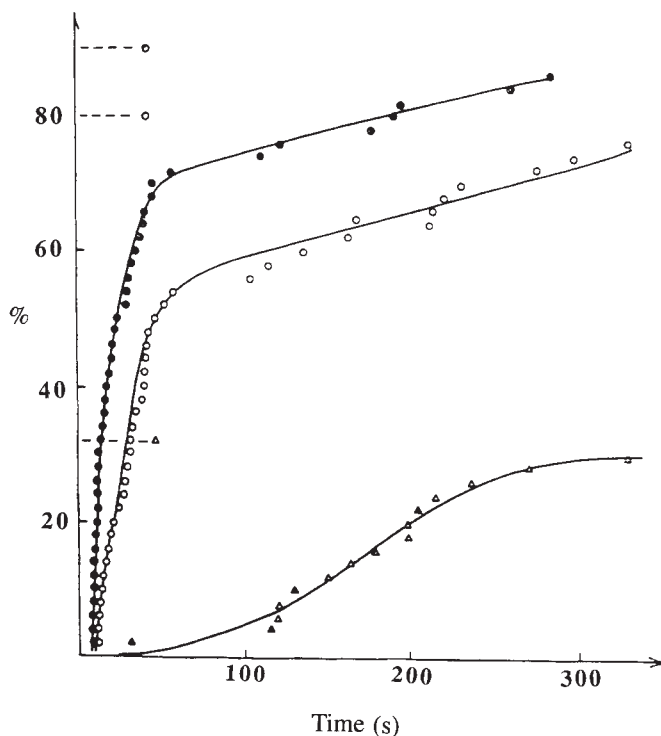


FIG. 6 Percentage of samples reacted as a function of time in the chlorite-thiosulphate reaction for three different stirring rates³⁰. Volume, 4.00 cm³; temperature, 30.0 °C, initial concentration of thiosulphate [Na₂S₂O₃]₀, 0.100 M; [NaClO₂]₀, 0.300 M; [NaOH]₀, 0.100 M. Stirring rate: filled circles (top curve), 500 r.p.m.; open circles (middle curve), 620 r.p.m.; triangles (bottom curve), 700 r.p.m. Note that successive points are equidistant vertically, being separated by 2%, as each distribution is based on 50 measurements.

smoothly. In three of the curves, we observe a sharp drop in pH at approximately 3, 5 and 9 minutes; in the other two, this decrease occurs at times greater than 20 minutes. When an acid-base indicator like phenolphthalein is added to the solution, the pH change is accompanied by a striking colour change. In an 'ordinary' clock reaction, or this same reaction in a different range of reactant concentrations, the sharp pH drop would occur at the same time, to within a few seconds, in all of the curves. What, other than experimental error, might lead to such irreproducibility, sometimes as much as several orders of magnitude, in the reaction times of the anomalous experiments?

Careful efforts to remove all sources of variability among experiments³⁰ met with total failure. Despite elaborate schemes to ensure that all experiments were the same with regard to temperature, initial concentrations, exposure to light, vessel surface, age of solutions and mixing procedure, the reaction times still varied over a wide range. The distribution of reaction times for, say, a set of 100 replicate experiments under the same conditions was, however, statistically reproducible. If the conditions such as reactant concentrations, temperature, stirring rate or, most remarkably, the volume of the vessel, were varied, the distribution of reaction times changed significantly. In Fig. 6, I show a cumulative probability distribution of reaction times for three sets of experiments carried out with three different stirring rates. In these experiments, both the mean and the standard deviation of the reaction-time distribution increase as the stirring rate increases or as the reaction volume decreases.

Mechanistic analysis of the chlorite-thiosulphate reaction as well as consideration of the reaction-time distributions led to the following qualitative explanation for the observed behaviour. The pH rise and subsequent drop result from competition between two reaction pathways starting from chlorite and thiosulphate:



Reaction (3) is responsible for the initial pH rise; its rate is proportional to $[\text{H}^+]$. Reaction (4) accounts for the pH drop. It is 'supercatalytic', with a rate proportional to $[\text{H}^+]^2[\text{Cl}^-]$. In the bulk of the solution, reaction (3) is the dominant pathway. However, if in some small region a fluctuation leads to a sufficiently high concentration of hydrogen ions, the autocatalysis in reaction (4) can produce a rapid local build-up of $[\text{H}^+]$, which spreads through the solution, causing the observed rapid pH drop.

The stochastic nature of the reaction-time distribution results from the fact that these supercritical concentration fluctuations are rare, random events. The observed effects of stirring and volume on the distribution support this interpretation. By increasing the effectiveness of the mixing, one decreases the likelihood that any given fluctuation will grow to critical size before it is dissipated by mixing with the high-pH bulk. Note that if the stirring is stopped at any time after the first few seconds, all samples react completely within three seconds. The more puzzling volume effect can be understood simply in terms of the number of possible microvolumes within the sample in which a supercritical fluctuation can occur. The larger the volume of the system, the more likely such a fluctuation and the shorter the mean reaction time.

Although the behaviour described above is rare, it is by no means unique to the chlorite-thiosulphate reaction. Similar distributions of reaction times are found in the chlorite-iodide reaction³¹ in a concentration range where it shows high-order autocatalysis. Ferrone *et al.*³² and Hofrichter³³ have observed a stochastic variation in the distribution of times at which polymerizing sickle haemoglobin molecules undergo a sharp transition from the monomeric to the polymeric state. They suggest a route involving a highly autocatalytic "double nucleation" mechanism that shares important features with the scheme out-

lined above for the chlorite-thiosulphate reaction and with Kondepudi's description of secondary nucleation in sodium chlorate crystallization. The primary nucleus is a polymer of sickle haemoglobin that exceeds a critical length, beyond which it can serve as a template for the formation of new polymer chains.

Lemarchand *et al.*³⁴ reported a stochastic distribution of ignition times in experiments on the kinetics of cool flames in mixtures of *n*-pentane and air. Combustion reactions tend to be poorly mixed, because of the complex spatial structure of flames, and autocatalytic, both because of the occurrence of radical chain reactions and because the reaction rates increase as the temperature rises.

At least two theoretical approaches provide a more quantitative account of these stochastic reaction-time distributions. Szabo³⁵ has used the theory of first passage times and a simple birth and death model to obtain an analytical expression that is in good agreement with the sickle haemoglobin polymerization results. Sagués *et al.*³⁶ use a more detailed chemical model with noise terms incorporated through the kinetic constants to simulate the reaction-time distributions in experiments on the chlorite-iodide reaction. In both cases, mixing efficiency serves as a parameter: in the Szabo approach it determines the initial nucleation probability; in the noise-term method the intensity of the noise is empirically correlated with the stirring rate.

Conclusion

The past decade has witnessed a growing understanding of the rich variety of behaviour of which nonlinear, and especially autocatalytic, chemical reactions are capable. Typically, more complex behaviour emerges when a control parameter such as the temperature or a reactant concentration is varied past a critical value. The studies described here suggest that mixing effectiveness, difficult though it may be to quantify, can serve as an additional, more subtle control parameter. In addition, the interaction of mixing with autocatalysis can give rise to new phenomena such as chiral symmetry-breaking and stochastic behaviour. As we have seen, the interaction of autocatalysis and mixing gives rise to interesting effects in both open and closed systems.

From a theoretical point of view, the extreme cases of perfectly mixed, homogeneous systems and of unmixed, diffusive systems are far better understood than the intermediate case of partially mixed systems, though the latter are far more common in nature. An accurate description of the mixing process is still far beyond our current theoretical capabilities, but, as some of our examples show, even a rather crude model of the mixing when coupled to a reasonable chemical scheme can yield insight into many of the phenomena of interest.

Further study of the behaviour that can arise in imperfectly mixed autocatalytic systems should prove rewarding. There may well be new phenomena to be discovered. The notion of tuning the mixing efficiency to control the composition of the output from an industrial reactor is an appealing potential extension of the experiments that produce chirally pure sodium chlorate crystals. Workers at duPont (Roelofs and Wasserman, unpublished observations) have studied the oxidation of toluene in an effort to gain insight into the analogous *p*-xylene oxidation reaction that generates the starting material for polyester fibre. The reaction is autocatalytic and oscillatory. In a large reactor similar to those employed industrially, the phase and amplitude of the oscillations vary from place to place in the reactor. This variation can be reduced, although not totally eliminated, by increasing the rate at which the solution is stirred. The consequences of this spatial inhomogeneity for the yield and purity of the product are not yet known, but seem worth exploring further.

Finally, the possible consequences of imperfect mixing in living systems merit additional scrutiny. As noted at the outset, population growth is an inherently autocatalytic process. In a

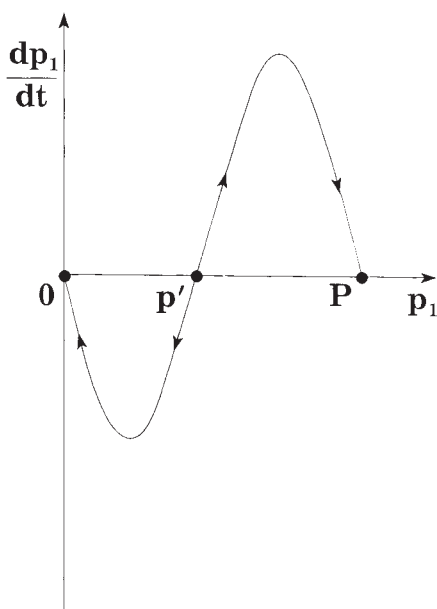
BOX 1 A simple population-dynamics example

Consider a set of species with populations $p_1, p_2, \dots$. Assume that each grows and dies at rates proportional to its population,

$$dp_i/dt = (b_i - d_i)p_i = r_i p_i \quad (5)$$

where b_i and d_i are the *per capita* birth and death rates, respectively, and r_i is the net *per capita* growth rate of the i th species. If the food supply is unlimited and if there are no (for example, predator-prey) interactions between species, then all populations grow exponentially, and the species with the largest growth rate r_i will eventually come to dominate the population regardless of the initial populations p_{i0} . A similar conclusion holds for any other growth law if the growth is unconstrained.

In some cases, for example that of equation (1), one may wish to take the effects of a limited food supply or other environmental constraints into account by making the growth rate a nonlinear function of the population. Alternatively, one might include these effects by constraining the total population. We can maintain the total population, $P = \sum p_i$, constant by adding a flux term to our population growth equation. The particular form of the constraint used in equations (6a) and (6b) below for a pair of competing species is referred to by Eigen and Schuster³⁷ in their models of molecular evolution as



the constraint of constant organization. With initial populations p_{10} and $p_{20} = P - p_{10}$, the growth equations are

$$dp_1/dt = r_1 p_1 - (p_1/P)(r_1 p_1 + r_2 p_2) \quad (6a)$$

$$dp_2/dt = r_2 p_2 - (p_2/P)(r_2 p_2 + r_1 p_1) \quad (6b)$$

Note that $dp_1/dt + dp_2/dt = 0$. Because $p_2 = P - p_1$, we can eliminate p_2 from equation (6a) to obtain a single growth equation for p_1 :

$$dp_1/dt = r_1 p_1 - (p_1/P)[r_1 p_1 + r_2 (P - p_1)] \\ = [(r_2 - r_1)/P][p_1 (P - p_1)] \quad (7)$$

Equation (7) can be solved exactly for $p_1(t)$ to yield

$$p_1(t) = P/[1 - (1 - P/p_{10}) \exp [(r_2 - r_1)t]] \quad (8)$$

As $t \rightarrow \infty$ our expression for $p_1(t)$ approaches 0 if $r_2 > r_1$; it approaches P if $r_2 < r_1$. We thus have all-or-none selection of the faster growing species, independent of the initial populations, so long as neither starts at zero.

Now consider a pair of species evolving with a nonlinear growth rate, say quadratic, again under the constraint of constant organization. The evolution equations for this system take the form

$$dp_1/dt = r_1 p_1^2 - (p_1/P)(r_1 p_1^2 + r_2 p_2^2) \quad (9a)$$

$$dp_2/dt = r_2 p_2^2 - (p_2/P)(r_1 p_1^2 + r_2 p_2^2) \quad (9b)$$

Substitution of $P - p_1$ for p_2 in equation (9a) yields, after some algebraic manipulation,

$$dp_1/dt = -[(r_1 + r_2)/P]p_1[p_1 - r_2 P/(r_1 + r_2)](p_1 - P) \quad (10)$$

Equation (10) can be solved exactly for $p_1(t)$, but a graphical solution is more revealing. In the box figure, I have plotted dp_1/dt versus p_1 : the arrows show the direction in which the system evolves. From the sign of the time derivative, we see that if p_1 is initially less than $p' = r_2 P/(r_1 + r_2)$ it will decrease until it reaches a final value of zero, that is, species 1 will die out. If $p_1 > p'$, then it will grow until it reaches P . We again have all-or-none selection, but now the winner of the battle depends, not simply on the growth rate, but on whether or not p_1 initially exceeds p' . Survival of the fittest is no longer the operative principle; it matters who gets in first. Note that

$$p' = r_2 P/(r_1 + r_2) = r_2(p_1 + p_2)/(r_1 + r_2)$$

so that the condition $p_1 > p'$ is equivalent to $p_1(r_1 + r_2) > r_2(p_1 + p_2)$, or equivalently, to $r_1 p_1 > r_2 p_2$. Selection that is dependent on the initial conditions occurs for any growth law that exceeds linearity in its dependence on the populations. Similar conclusions can be reached for an arbitrary number of species³⁷.

perfectly mixed system of species competing for a common, fixed food supply, there will be all-or-none selection (see Box 1). The situation resembles somewhat the L- and D-crystals competing for the dissolved sodium chlorate in the experiments of Kondrupi *et al.*²⁶.

The "fittest" species will ultimately win out as shown in Box 1. The definition of fitness varies with the rate law for reproduction and the environment in which the competition takes place. If the food supply is unlimited or if the growth rate is linear, even a single reproductively capable individual will give rise to descendants who will dominate an established population of lower fecundity (or survivability). If the food supply is limited and the growth rate is nonlinear, then history plays a role. In order to win the competition, a new species must first reach a certain population threshold at which its absolute, not its relative rate of reproduction is the largest in the system. For species with quadratic growth rates r_1 and r_2 and populations p_1 and p_2 , this means that species 1 will be the final survivor if and only if p_1 initially exceeds $(r_2/r_1)p_2$. In effect, a favourable mutation cannot survive unless there is a way for its population to reach a critical level before it is swamped by competition from its less able but more populous neighbours.

One way in which real entities avoid the wastefulness of

all-or-none selection is through less-than-perfect mixing. Mobile organisms can isolate themselves from direct competition with their better adapted fellows by spatially segregating themselves, by choosing a niche. Darwin recognized that more exotic, though perhaps 'less fit' species were likely to be found in isolated habitats such as islands. Eigen and Schuster³⁷ have considered a model of the earliest self-reproducing macromolecules forming from their monomeric components in a puddle of prebiotic broth. They find that, to the extent that different molecules can find different puddles in which to grow, or that compartmentalization can arise from the development of primitive cell membranes, the mixture will be able to support a richer variety of complex species. Kareiva and Wennergren³⁸ have recently reviewed the fascinating literature on how landscape patterns and species mobility affect population dynamics in complex ecological systems. Again, the rates at which species mix and spread are found to have dramatic effects on the behaviour of the system.

Whether the sorts of experiments described here will give us further insight into controlling the outputs of vats of industrially important chemicals or will allow us to understand better the mysteries of molecular or species evolution remains to be seen. What is already clear is that the interaction of autocatalytic

processes with imperfect mixing leads to a rich variety of phenomena and that there is still much to be learned about the behaviour of such systems. □

Irving R. Epstein is at the Department of Chemistry and Volen Center for Complex Systems, Brandeis University, Waltham, Massachusetts 02254-9110, USA.

1. May, R. M. *Nature* **261**, 459–467 (1976).
2. Jahnke, W. & Winfree, A. T. *J. chem. Educ.* **68**, 320–324 (1991).
3. Devreotes, P., Fontana, D., Klein, P., Sherring, J. & Theibert, A. *Meth. Cell. Biol.* **28**, 299–331 (1987).
4. Lechleiter, J., Girard, S. E., Peralta, E. G. & Clapham, D. *Science* **252**, 123–126 (1991).
5. Turing, A. M. *Phil. Trans. R. Soc. B* **237**, 37–72 (1952).
6. Castets, V., Dulos, E., Boissonade, J. & De Kepper, P. *Phys. Rev. Lett.* **64**, 2953–2956 (1990).
7. Malthus, T. R. *An Essay on the Principle of Population* 14 (Oxford Univ. Press, London, 1798).
8. Epstein, I. R. *J. chem. Educ.* **66**, 191–195 (1989).
9. Villermaux, J. *Rev. chem. Engng.* **7**, 51–108 (1991).
10. Roux, J. C., De Kepper, P. & Boissonade, J. *J. Phys.* **97A**, L168–L170 (1983).
11. De Kepper, P., Boissonade, J. & Epstein, I. R. *J. phys. Chem.* **94**, 6525–6536 (1990).
12. Luo, Y. & Epstein, I. R. *J. chem. Phys.* **85**, 5733–5740 (1986).
13. Kumpinsky, E. & Epstein, I. R. *J. chem. Phys.* **82**, 53–57 (1985).

14. Menzinger, M. & Dutt, A. K. *J. phys. Chem.* **94**, 4510–4514 (1990).
15. Schmitz, R. A., Graziani, K. R. & Hudson, J. L. *J. chem. Phys.* **67**, 3040–3044 (1977).
16. Roux, J. C., Rossi, A., Bachelart, S. & Vidal, C. *Phys. Lett.* **A77**, 391–393 (1980).
17. Roux, J. C., Simoyi, R. H. & Swinney, H. L. *Physica D* **8**, 257–266 (1983).
18. Györgyi, L. & Field, R. J. *J. chem. Phys.* **91**, 6131–6141 (1989).
19. Field, R. J. & Noyes, R. M. *J. chem. Phys.* **60**, 1877 (1974).
20. Györgyi, L. & Field, R. J. *J. phys. Chem.* **95**, 6594–6602 (1991).
21. Györgyi, L. & Field, R. J. *Nature* **355**, 808–810 (1992).
22. Maelko, J. & Epstein, I. R. *J. chem. Phys.* **80**, 3175–3178 (1984).
23. Alamgir, M. & Epstein, I. R. *Int. J. chem. Kinet.* **17**, 429–439 (1985).
24. Olsen, L. F. & Degn, H. *Nature* **267**, 177–178 (1977).
25. Aguda, B. D. & Larter, R. J. *Am. chem. Soc.* **113**, 7913–7916 (1991).
26. Kondepudi, D. K., Kaufman, R. J. & Singh, N. *Science* **250**, 975–976 (1990).
27. Kondepudi, D. K. & Sabanayagam, C. *Chem. Phys. Lett.* **217**, 364–368 (1994).
28. Metcalfe, G. & Ottino, J. M. *Phys. Rev. Lett.* **72**, 2875–2878 (1994).
29. Swanson, P. D. & Ottino, J. M. *J. Fluid Mech.* **213**, 227–249 (1990).
30. Nagypál, I. & Epstein, I. R. *J. phys. Chem.* **90**, 6285–6292 (1986).
31. Nagypál, I. & Epstein, I. R. *J. chem. Phys.* **89**, 6925–6928 (1988).
32. Ferrone, F. A., Hofrichter, J. & Eaton, W. A. *J. molec. Biol.* **183**, 591–610, 611–631 (1985).
33. Hofrichter, J. *J. molec. Biol.* **189**, 533–571 (1986).
34. Lemarchand, A., Ben Alm, R. I. & Nicolis, G. *Chem. Phys. Lett.* **162**, 92–98 (1989).
35. Szabo, A. *J. molec. Biol.* **199**, 539–542 (1988).
36. Sagüés, F., Ramirez-Piscina, L. & Sancho, J. M. *J. chem. Phys.* **92**, 4786–4792 (1990).
37. Eigen, M. & Schuster, P. *The Hypercycle. A Principle of Natural Self-Organization* (Springer, Berlin, 1979).
38. Kareiva, P. & Wennergren, U. *Nature* (in the press).

ACKNOWLEDGEMENTS. I thank K. Kustin, R. Field and J. Pojman for helpful discussions. This work was supported by the US NSF.

Structural basis of cell–cell adhesion by cadherins

Lawrence Shapiro^{*}, Allison M. Fannon[†], Peter D. Kwong^{*}, Andrew Thompson[‡], Mogens S. Lehmann^{§||}, Gerhard Grübel[§], Jean-François Legrand^{§||}, Jens Als-Nielsen^{§||}, David R. Colman[†] & Wayne A. Hendrickson^{*#}

^{*} Department of Biochemistry and Molecular Biophysics, and [#] Howard Hughes Medical Institute, Columbia University, New York, New York 10032, USA

[†] Brookdale Center for Molecular Biology, Mount Sinai School of Medicine, New York, New York 10029, USA

[‡] European Molecular Biology Laboratory, BP 156 X, 38042 Grenoble, France

[§] European Synchrotron Radiation Facility, BP 220, 38043 Grenoble, France

^{||} Institut Laue Langevin, BP 156 X, 38042 Grenoble, France

Crystal structures of the amino-terminal domain of N-cadherin provide a picture at the atomic level of a specific adhesive contact between cells. A repeated set of dimer interfaces is common to the structure in three lattices. These interactions combine to form a linear zipper of molecules that mirrors the linear structure of the intracellular filaments with which cadherins associate. This cell-adhesion zipper may provide a mechanism to marshal individual molecular adhesive interactions into strong bonds between cells.

MANY cells of multicellular organisms are able to distinguish different cell types, generally adhering preferentially to cells of their own type^{1,2}. This selective cell adhesion, which arises as the result of expression of specific adhesive molecules on the cell surface, is fundamental to controlling the development and maintenance of tissues. The cadherin family of cell–cell adhesion receptors, whose members are expressed widely in the solid tissues of both vertebrates and invertebrates³, is central to this process⁴. N-cadherin^{5,6}, which first appears in the vertebrate nervous system during formation of the neural tube, functions in

morphogenesis and homeostasis in the nervous system and other tissues⁷.

Cadherins are transmembrane glycoproteins characterized by a distinctive sequence motif which is tandemly repeated in their extracellular segments⁵. We show here that these repeated sequences form autonomously folded protein modules ('cadherin domains'). Each cadherin subtype has a unique, usually homophilic, binding specificity that confers specific cell-adhesive properties on the cells in which it is expressed⁸. Experiments with molecular chimaeras⁶, neutralizing antibodies⁹ and peptide inhibitors¹⁰ suggest that the binding specificity resides in the first (N-terminal) cadherin domain.

The adhesive action of cadherins is dependent on the presence of extracellular Ca²⁺ ions¹¹, and it has long been known that Ca²⁺ affords substantial protection against exogenous proteases¹². Intracellular domains of cadherins mediate connections

* Permanent addresses: Laboratoire de Spectrométrie Physique (URA CNRS no. 8), Université J. Fourier-Grenoble 1, BP 87, 38402 Saint Martin d'Hères Cedex, France (J.-F.L.); Risø National Laboratory, DK 4000 Roskilde, Denmark (J.A.-N.).

TABLE 1 Statistics for data collection, phase determination and refinement

(a) Undulator MAD experiments

Data Collection† (20 to 3.0 Å)					Four-wavelength experiment*					
Wavelength (Å)		Reflections (N)	Completeness (%)	$\langle I/\sigma(I) \rangle$			R_{sym} (%)			
1.4013 (pre-edge)		4,811	90.0	10.5			6.1			
1.3852 (peak)		4,771	89.2	10.6			6.3			
1.3857 (edge)		4,762	89.1	10.2			6.5			
1.3847 (remote)		4,793	89.6	11.3			6.0			
MAD structure factor ratios‡										
λ (Å)		20.0 < d < 4.0 Å			4.0 < d < 3.0 Å				Scattering factors (e)	
	1.4013	1.3857	1.3852	1.3847	1.4013	1.3857	1.3852	1.3847	f'	f''
1.4013	0.084 (0.076)	0.067	0.051	0.044	0.114 (0.111)	0.089	0.086	0.077	−12.48	3.80
1.3857		0.110 (0.069)	0.049	0.067		0.140 (0.127)	0.085	0.089	−31.22	17.20
1.3852			0.149 (0.072)	0.039			0.155 (0.119)	0.082	−13.97	29.17
1.3847				0.102 (0.071)				0.124 (0.120)	−6.67	17.34

Data collection† (20 to 3.0 Å)					Five-wavelength experiment*				
Wavelength (Å)	Reflections (N)	Completeness (%)	$\langle I/\sigma(I) \rangle$			R_{sym} (%)		Undulator gap (mm)	
1.3857 (ascending edge)	4,480	83.8	8.2			6.9		21.6	
1.3852 (peak)	4,450	83.2	7.7			7.5		21.6	
1.3847 (descending edge)	4,359	81.5	7.4			7.6		21.6	
1.3784 (remote 1)	4,394	82.2	7.0			7.8		21.7	
1.2862 (remote 2)	4,448	83.2	9.3			6.4		22.5	

MAD structure factor ratios‡									
λ (Å)	20.0 < d < 4.0 Å				4.0 < d < 3.0 Å				Scattering factors (e)
	1.3857	1.3852	1.3847	1.3784	1.2862	1.3857	1.3852	1.3847	f' f''
1.3857	0.075 (0.027)	0.041	0.060	0.057	0.072	0.078 (0.075)	0.059	0.067	-28.33 14.84
1.3852		0.105 (0.032)	0.036	0.044	0.065		0.101 (0.088)	0.066	-21.50 30.23
1.3847			0.072 (0.031)	0.040	0.059			0.097 (0.074)	-10.71 20.35
1.3784				0.059 (0.032)	0.059			0.114 (0.089)	-14.45 11.84
1.2862					0.051 (0.028)				-9.03 9.01

MAD phasing§

$$R(|^{\circ}F_T|) = 0.063 \quad \langle \Delta(\Delta\phi) \rangle = 58.5^{\circ} \quad \langle m \rangle = 0.88$$

$$R(|^{\circ}F_A|) = 0.451 \quad \langle \sigma(\Delta\phi) \rangle = 20.3^{\circ}$$

(b) UO₂ MAD experiment

Data collection† (15.0 to 1.9 Å)												
Wavelength (Å)		Reflections (N)		Completeness (%)				$\langle I/\sigma(I) \rangle$		R_{sym} (%)		
0.7263 (pre-edge)		33,410		96.6 (92.9)				15.2		3.6		
0.7251 (edge)		33,144		95.6 (92.2)				14.3		3.6		
0.7284 (peak)		33,075		95.7 (92.0)				14.0		3.8		
0.7246 (descending edge)		33,142		95.8 (92.3)				13.9		3.7		
0.7217 (remote)		33,036		95.5 (91.2)				14.0		3.8		
MAD structure factor ratios‡												
λ (Å)	15.0 < d < 3.0 Å					3.0 < d < 1.9 Å					Scattering factors (e)	
	0.7263	0.7251	0.7248	0.7246	0.7217	0.7263	0.7251	0.7248	0.7246	0.7217	f'	f''
0.7263	0.035 (0.026)	0.028	0.023	0.025	0.026	0.060 (0.050)	0.056	0.055	0.053	0.056	−21.10	4.08
0.7251		0.060 (0.026)	0.029	0.031	0.035		0.089 (0.050)	0.054	0.058	0.063	−34.72	7.92
0.7248			0.075 (0.030)	0.023	0.027			0.104 (0.057)	0.052	0.057	−24.87	10.30
0.7246				0.069 (0.026)	0.024				0.098 (0.052)	0.054	−17.43	9.62
0.7217					0.060 (0.028)					0.089 (0.060)	−13.26	8.40

MAD phasing§

$$R(|^{\circ}F_T|) = 0.051 \quad \langle \Delta(\Delta\phi) \rangle = 36.8^{\circ} \quad \langle m \rangle = 0.93$$

$$R(|^{\circ}F_A|) = 0.419 \quad \langle \sigma(\Delta\phi) \rangle = 18.2^{\circ}$$

TABLE 1—continued

(c) Refinement

	Resolution limit (Å)	Temp. (°C)	Refinement data sets¶						<i>R</i> _{sym} (622) (%)	<i>R</i> _{sym} (321) (%)
			Reflections (N)		Completeness (%)		<i>I</i> /σ(<i>I</i>)			
			(622)	(321)	(622)	(321)				
Crystal type A:	1.9	−160	18,803 (10,734)	33,961 (19,763)	90.0 (93.4)	85.3 (94.2)	13.2	4.5	4.1	
Crystal type B:	2.1	20	(7,986)	(15,198)	(93.3)	(93.2)	11.5	11.6	4.7	
Crystal type C:	1.9	−160	(Peak wavelength data set from UO ₂ MAD experiment)							
Refinement statistics¶										
Crystal type A (P6 ₃ 22) model: 1 molecule per asymmetric unit, 99 residues, 202 waters, one Yb ³⁺ ion (973 atoms)										
Crystal type B (P321) model: 2 molecules per asymmetric unit, 199 residues, 231 waters, two Yb ³⁺ ions (1,761 atoms)										
Crystal type C (P2 ₁ 2 ₁ 2) model: 2 molecules per asymmetric unit, 197 residues, 316 waters, one UO ₂ ²⁺ ion (1,875 atoms)										
	<i>d</i> -spacings (Å)	Reflections (N)	<i>R</i> -value (%)	⟨ <i>B</i> -value⟩ (Å ²)	Bonds (Å)	R.m.s. deviations Angles (Å) <i>B</i> -values (Å ²)				
Crystal type A:	5.0–1.9	17,546	21.7	25.6	0.015	2.1	1.9			
Crystal type B:	5.0–2.1	13,950	22.4	32.6	0.012	1.8	2.3			
Crystal type C:	5.0–1.9	31,010	19.5	17.1	0.013	1.8	1.2			

Cloning and expression. NCD1 was produced as a glutathione S-transferase (GST) fusion protein in *E. coli*. A cDNA encoding residues 1–108 of the mature protein was produced by reverse transcription and PCR from a mouse brain mRNA library. This cDNA was used as a template for PCR reactions whose products were subcloned into the pGEX-2T (ref. 39) expression vector, which produces a thrombin cleavage site. The resulting construct coded for two extra N-terminal residues (Gly-Ser) and proved ultimately to have a PCR mutation (D27G). **Purification and crystallization.** The fusion protein was isolated by affinity chromatography on glutathione-agarose. NCD1 was then isolated by cleavage from GST with thrombin, and subsequent purification by ion exchange chromatography on a mono-Q column. This material was stored frozen in 1 mM Tris, pH 8.0, at a concentration of 35 mg ml^{−1} and was used directly for crystallization. Crystals were produced by the vapour diffusion method at 19.5 °C in hanging drops containing 2–3 µl protein solution and 1 µl of the reservoir solution of 0.8 M Li₂SO₄, 0.1 M Na acetate, pH 5.0, 20 mM Yb acetate. The crystals are large hexagonal rods with dimensions typically 0.2 × 0.2 × 1.0 mm. At 20 °C (crystal type B and undulator MAD experiment) the crystals have unit-cell dimensions $a = b = 79.4$ Å, $c = 74.9$ Å. Crystal type A has unit-cell dimensions $a = b = 77.6$ Å, $c = 74.9$ Å when frozen at −160 °C. R_{sym} values for data sets of many different NCD1 crystals acquired on a rotating anode X-ray source (data not shown) suggest that either crystal type can exist at room temperature, and that some crystals may have mixtures of the two lattices. Only type A crystals have been observed in the frozen state. The UO₂ crystals were produced by the vapour diffusion method at 19.5 °C in hanging drops containing 2 µl protein solution at 35 mg ml^{−1} and 1 µl of the reservoir solution of 0.5–0.6 M (NH₄)₂SO₄, 0.1 M Na acetate, pH 4.0, 5 mM UO₂ acetate. The crystals grow as plates up to 0.4 × 0.4 × 0.2 mm, and diffract strongly to Bragg spacings up to 1.9 Å. The crystals are in space group P2₁2₁2, with two molecules per asymmetric unit and unit-cell dimensions $a = 75.2$ Å, $b = 57.5$ Å, $c = 52.5$ Å. A single UO₂²⁺ ion is coordinated quasi-symmetrically between analogous sites of two protomers. **Undulator MAD data collection and processing.** MAD data were collected on the TROIKA beamline⁴⁰ at the ESRF using a MAR Research imaging plate (IP) system. The central cone (0.5 × 0.5 mm) of the undulator beam produced at 6 GeV and 100 mA was monochromatized by a thin, water-cooled Si 220 crystal⁴¹ and deflected, without focusing, by a SiC mirror for harmonic rejection. A 0.3 mm collimator was used. Oscillation ranges of 2.5° to 3.0° were used with exposure times of 90–120 s. Bijvoet mates were collected on the same or adjacent imaging plate, assuming the mirror symmetry of Laue group 622. IP data were reduced to integrated intensities with the program DENZO⁴². ROTAVATA⁴³ was used to calculate scaling parameters for each IP, which were applied with a modified version of the program AGROVATA⁴³ which does not merge redundant reflections. Parameterized local scaling and phase evaluations were carried out with the MADSYS programs¹⁸. Residual systematic errors, which ultimately proved to be due to lattice pseudosymmetry, were further reduced by applying scale factors K for each reflection h across its set of wavelengths λ so as to meet the conditions $(|F(h)| - K|F(-h)|) = 0$ for centric reflections, and $(|F(h)| - K|F(-h)|) = Q''(\lambda)$ for general reflections. **Yb refinement data collection and processing.** Data for refinement were collected at beamline X4A of the National Synchrotron Light Source at Brookhaven National Laboratory. Oscillation ranges of 1.9° to 2.2° were used with exposure times of 90–120 s. Data were collected on Fuji IPs digitized with a Fuji scanner. DENZO was used to extract integrated intensities, and scaling was done with the CCP4 package of programs. P321 data were obtained from two crystals at 20 °C. Very few Bijvoet mates were taken for these crystals for space group P321 because at the time of data collection it was thought the crystals were in space group P6₃22. The P6₃22 refinement data were obtained from a single crystal frozen at −160 °C using paratone-n as a cryoprotectant. **UO₂ MAD data collection and processing.** Crystals were cryoprotected by soaking for 2 min each in 15% and then 30% glycerol plus the mother liquor of 0.5–0.6 M (NH₄)₂SO₄, 0.1 M Na acetate, pH 4.0, 5 mM UO₂-acetate. MAD data were collected from two frozen crystals at beamline X4A of the National Synchrotron Light Source at Brookhaven National Laboratory. Oscillation ranges of 1.5° were used with exposure times of 120–300 s. Bijvoet mates were collected on the same or adjacent imaging plate. These IP data were processed with DENZO and CCP4 scaling routines as for the undulator. Phases and figures of merit were extracted with the MADSYS programs. **Model building and refinement.** An α -carbon trace was built into the Yb-phased density using minimaps, and this trace was used to produce an initial model which was refined with one round of simulated annealing and several rounds of minimization and rebuilding. The R -value reached about 26% against a 2.5-Å data set combined from peak wavelength measurements in both undulator MAD experiments ($R_{\text{sym}} = 8.7\%$, with 90.5% completeness). Refinement was completed with a model derived from the UO₂ MAD experiment. For this experiment, an α -carbon trace was built into the MAD-phased density using the program TOM⁴⁴, and the LEGO subroutine of the program O (ref. 44) was used to produce an initial model. After refinement by least-squares minimization using the program XPLOR⁴⁵, this model was positioned into the undulator MAD-phased density and subsequently was refined against both the P6₃22 and P321 data sets. Refinement continued with an initial round of simulated annealing, and several rounds of least-squares refinement. The β -carbon and side chain of Asp 27 were not visible in the refined $2|F_o| - |F_c|$ electron density of any of the crystal forms. Resequencing of the expression vector showed that this residue had mutated to glycine. Subsequently, we determined the structure of a truncated native construct which terminated after Asp 100. This protein crystallized isomorphously with type C crystals under the same conditions. We found only minor changes in the structure around Asp 27.

* R_{sym} and completeness statistics shown for the undulator MAD experiment are given for space group P6₃22, in which phases were derived. R_{sym} values were typically about 1% lower in point group 321, for example 4.9% for the remote wavelength of the four-wavelength experiment. Completeness for these experiments in point group 321 is less than 50% because Bijvoet mates were not collected for this point group.

† Completeness statistics are calculated with Bijvoet mates considered inequivalent. $R_{\text{sym}} = 100 \times \sum_h \sum_l |I(h) - \langle I(h) \rangle| / \sum_h \sum_l I(h)$.

‡ Table values represent r.m.s. ($\Delta|F|$)/r.m.s. ($|F|$), where $\Delta|F|$ is the Bijvoet difference at one wavelength (diagonal elements) or the dispersive difference between two wavelengths (off-diagonal elements). The values in parentheses are the ratios for centric Bijvoet reflections, which would be equal to zero for perfect data; these serve as an estimate of the noise in the anomalous signals.

§ Value shown are for data between 20.0 and 3.0 Å. $R = \sum_h \sum_l |F_o(h) - \langle |F(h)| \rangle| / \sum_h \sum_l |F_o(h)|$. $F_o(h)$ is the structure factor due to normal scattering from all the atoms, $F_A(h)$ is the structure factor due to normal scattering from the anomalous scatterers only, and $\Delta\phi(h)$ is the phase difference between $F_o(h)$ and $F_A(h)$. $\Delta(\Delta\phi)$ is the difference between two independent determinations of $\Delta\phi(h)$. Values given are based on calculations that did not include reflections with $m = 0$; $\langle m \rangle$ is the mean figure of merit including reflections with $m = 0$.

|| Numbers in parentheses give statistics calculated without distinguishing Bijvoet mates.

¶ A subset of the data (5%) was excluded from the refinement and used for the free R -value calculation. All data with $|F| > 2\sigma$ were used for refinement.

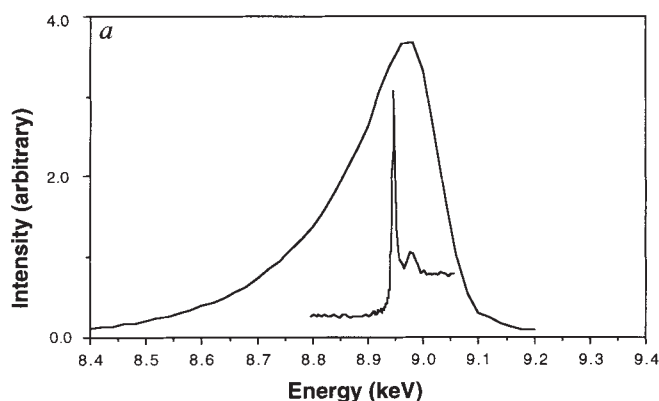
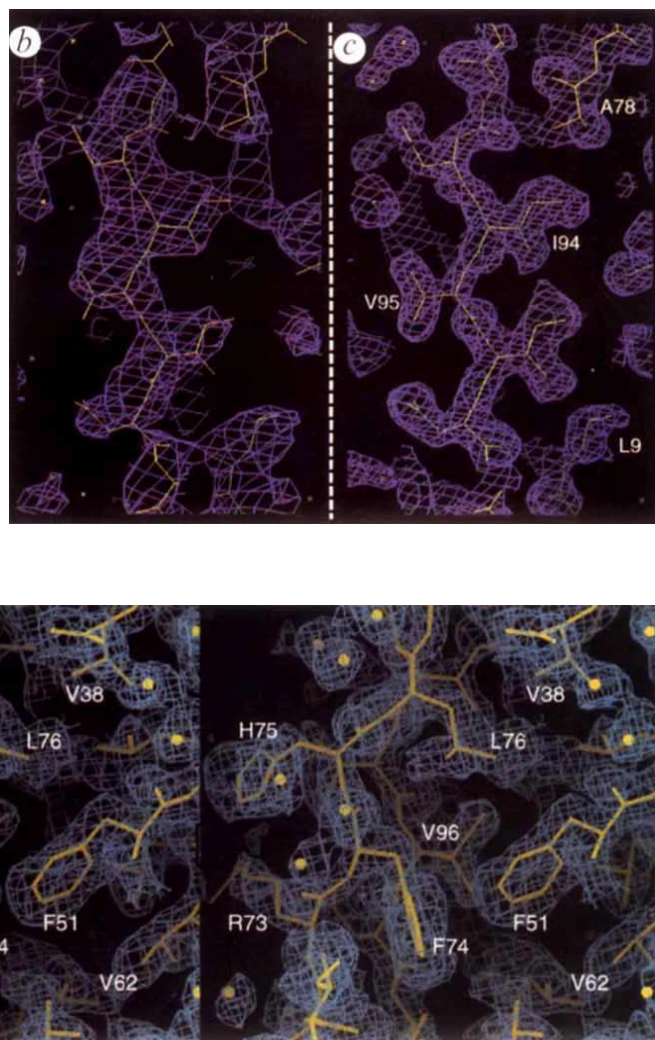


FIG. 1 MAD phasing experiments. *a*, X-ray fluorescence spectrum of the Yb^{3+} L_{III} edge transition of an NCD1 crystal superimposed on an energy scan of the third harmonic of the TROIKA undulator tuned to this edge with an undulator gap of 21.6 mm, as was used in the four-wavelength MAD experiment. *b*, Electron density distribution of region of the 10.0 to 3.0 Å undulator MAD-phased map encompassing a portion of the G strand, shown with the final $P6_322$ model of NCD1. *c*, The same region is shown with a map calculated with $2|F_o| - |F_c|$ coefficients for data between 5.0 and 1.9 Å Bragg spacings. *d*, Stereo diagram of a region of the experimental electron density map produced by the five-wavelength uranium MAD experiment. Superposed is the refined atomic model. This map was made with data that included reflections from 10.0 to 1.9 Å Bragg spacings. Each electron density map is contoured at 1.0σ .



between cadherins and cytoskeletal filaments^{13,14}, and it has been shown that cadherin adhesion is also dependent on the function of the intracellular domain^{15,16}.

Here we describe high-resolution crystal structures of the first cadherin domain of murine N-cadherin (NCD1) in three different crystal lattices. All of these structures have two dimer interfaces that combine to form a zipper-like supermolecular ribbon which we believe to be the fundamental unit of cadherin adhesive function. A single Ca^{2+} -binding site is located far from the adhesive interface, at the prospective interface between successive cadherin domains. We propose a model for connections between these domains that are rigidified by calcium. In our determination of the initial structure, we used undulator radiation for the first time in a multiwavelength anomalous diffraction (MAD) experiment.

Undulator MAD experiments

Undulator insertion devices at third-generation synchrotron radiation facilities generate extraordinarily bright X-ray beams¹⁷, and should extend the applicability of crystallographic analysis to previously intractable problems. We attempted here to adapt this new radiation source to the MAD phasing technique¹⁸. MAD requires tunable radiation, and previous experiments have been performed on bending magnets and wigglers that have continuous emission spectra. In contrast, undulator radiation is concentrated into narrow ($\sim 2\%$ bandwidth) harmonic peaks. Their position in energy can, however, be adjusted by tuning the magnetic gap of the undulator, and by so doing we performed two

MAD experiments at the TROIKA undulator beamline at the ESRF¹⁷.

Type A crystals were grown from recombinant fragment NCD1 complexed with Yb^{3+} as a Ca^{2+} replacement. Using these crystals, we tried two different protocols to circumvent the restricted energy range of the undulator harmonic peak. In the first experiment, we exploited the sharp transition of the Yb L_{III} edge and performed an entire MAD experiment with wavelengths chosen to fall within the third harmonic of the undulator tuned to the appropriate energy (Fig. 1*a*). Owing to the intrinsic undulator brilliance, the Yb absorption spectrum and associated anomalous scattering factors are limited only by natural line-widths of the transition. In the second experiment, the position of the undulator harmonic peak was tuned to track the MAD experiment by recurrently varying the magnet gap of the insertion device. This second approach allowed collection of data over a greater energy range to enable wavelengths with larger f' differences to be used in phase calculations.

The undulator data collection was successful in both experiments (Table 1*a*); however, the calculation of accurate phases was hampered by two unexpected factors related to the protein crystals. First, subsequent analysis indicated that the several crystals used in data collection were a mixture of two different, but closely related crystal types. One lattice (type B, $P321$) is an isomorphous subgroup of the other (type A, $P6_322$) and the diffraction patterns are pseudosymmetrically equivalent, although distinct. Second, the Yb anomalous scatterer, from which MAD phasing was derived, was only poorly ordered, with

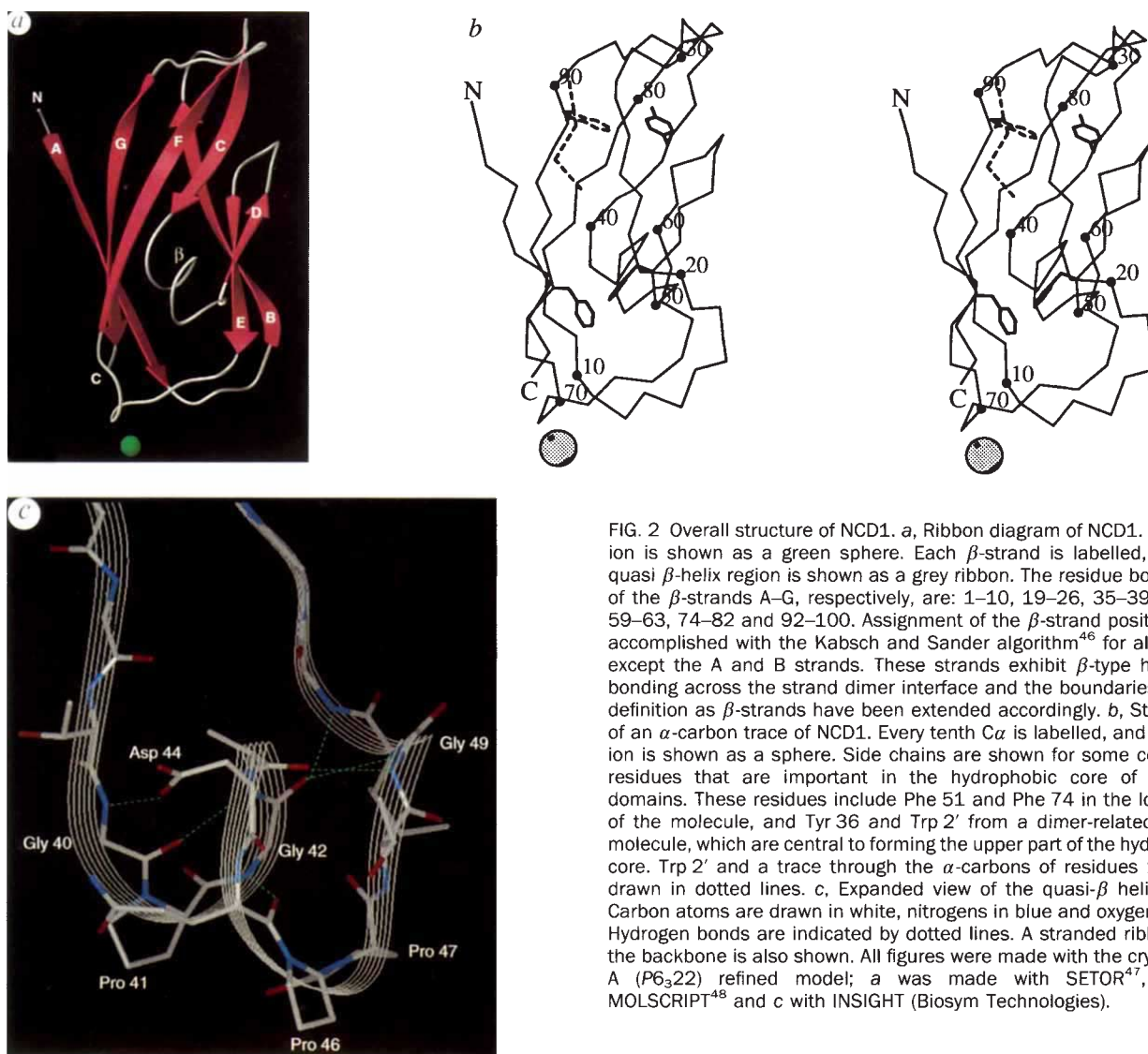


FIG. 2 Overall structure of NCD1. *a*, Ribbon diagram of NCD1. The Yb^{3+} ion is shown as a green sphere. Each β -strand is labelled, and the quasi β -helix region is shown as a grey ribbon. The residue boundaries of the β -strands A–G, respectively, are: 1–10, 19–26, 35–39, 51–54, 59–63, 74–82 and 92–100. Assignment of the β -strand positions was accomplished with the Kabsch and Sander algorithm⁴⁶ for all strands except the A and B strands. These strands exhibit β -type hydrogen-bonding across the strand dimer interface and the boundaries of their definition as β -strands have been extended accordingly. *b*, Stereo plot of an α -carbon trace of NCD1. Every tenth $\text{C}\alpha$ is labelled, and the Yb^{3+} ion is shown as a sphere. Side chains are shown for some conserved residues that are important in the hydrophobic core of cadherin domains. These residues include Phe 51 and Phe 74 in the lower part of the molecule, and Tyr 36 and Trp 2' from a dimer-related partner molecule, which are central to forming the upper part of the hydrophobic core. Trp 2' and a trace through the α -carbons of residues 1'–5' are drawn in dotted lines. *c*, Expanded view of the quasi- β helix region. Carbon atoms are drawn in white, nitrogens in blue and oxygens in red. Hydrogen bonds are indicated by dotted lines. A stranded ribbon fit to the backbone is also shown. All figures were made with the crystal type A ($P6_322$) refined model; *a* was made with SETOR⁴⁷, *b* with MOLSCRIPT⁴⁸ and *c* with INSIGHT (Biosym Technologies).

a B -factor of $\sim 60 \text{ \AA}^2$. Both of these factors limited the effective resolution of the experimental electron density maps to less than the nominal $3 \text{ \AA}$ resolution (Fig. 1*b*).

Structure determination

Despite the limitations from the crystals used in these undulator MAD experiments, a correct chain tracing was made and an atomic model constructed. While refinement of this model was underway, we independently determined the structure of NCD1 from crystals grown in the presence of uranyl acetate, which were found to be orthorhombic (type C, $P2_12_12$) with two NCD1 protomers per asymmetric unit (Table 1*b*). This crystal structure was also determined by the MAD method, but with bending magnet radiation. We measured data from two frozen crystals at five wavelengths near the uranium L_{III} edge, and obtained MAD phases for reflections to $1.9 \text{ \AA}$ Bragg spacings. The result was an experimental electron density map of extraordinary quality at $1.9 \text{ \AA}$ resolution (Fig. 1*d*), which we attribute to the combined advantages from MAD phasing and crystal freezing.

After high-resolution refinement of the type-C crystals, we returned to refinement of the type-A ($P6_322$) and type-B ($P321$) structures, starting from the C-lattice model, and completed the analysis with new high-angle type-A and B data sets obtained for each lattice type (Table 1). Comparison of these refined structures with the initial partially refined type A model shows that all side-chain orientations within the β -strands and quasi- β -helix region (see later) were modelled correctly, and that there were

no residues out of sequence register. The refined models for the type A structure at $1.9 \text{ \AA}$ resolution (Fig. 1*c*), the type B structure at $2.1 \text{ \AA}$ resolution, and the type C structure at $1.9 \text{ \AA}$ resolution, have crystallographic R -values of 21.7, 22.4 and 19.5%, respectively. Some of the residual error in the A and B structures may arise from cross-contamination of the two lattices. Differences between these two crystal types result from a hydrophobic zipper-like crystal lattice contact which slips a notch, separating the 6_3 axis of the $P6_322$ crystals into a crystallographic 3-fold axis and a near-perfect parallel molecular 2_1 axis, which is displaced by $1.06 \text{ \AA}$ from the 3-fold in the $P321$ crystals. Residues 101–108, which probably form part of the linker region between domains 1 and 2, are disordered in each model.

Structure of the protomer

The structure of cadherin domain NCD1 is shown in Fig. 2. It is a compact unit with overall dimensions of $\sim 45 \times 25 \times 25 \text{ \AA}$. Its polypeptide fold includes seven β -strands arranged as two β -sheets such that the amino and carboxy termini are at opposite ends of the domain. The Yb^{3+} ion is near the carboxy terminus. Successive β -strands are labelled A–G; strands D, E and B make up one sheet of the β -sandwich and strands A, G, F and C form the other sheet. The segment connecting strands C and D adopts an unusual, nearly helical structure which features a succession of β -turns and a set of β -like hydrogen bonds. We refer to this segment as a quasi- β -helix (Fig. 2*c*). Although it superficially resembles a helix, irregularities would preclude its classification

as a secondary structure motif. There is a turn of 3_{10} helix in the segment connecting strands B and C. This segment has high atomic mobility (α -carbon B -factors $> 30 \text{ \AA}^2$), and density is not observed for residues 30 and 31 in one of the type B protomers, nor for residues 27–32 in one of the type C protomers. Overall, the five independent NCD1 copies in the three crystal structures are very similar; in pairwise superpositions, r.m.s. deviations are 0.33–0.74 $\text{\AA}$ for $C\alpha$ atoms. The largest deviations are between the Yb^{3+} and UO_2^{2+} ion-binding loops.

Structure-based sequence alignments suggest to us that all five extracellular domains of vertebrate cadherins, and other cadherin family members as well, share a common folding topology. This topology is similar to that of immunoglobulin-like molecules, and indeed the NCD1 structure can be superimposed quite well on variable-like domains, despite vanishingly low sequence similarity and distinctive core-defining residues. We describe further details of the cadherin fold, its relationship to other molecules, and evolutionary implications elsewhere⁵⁰.

Strand-dimer interface

Each of the three crystal structures contains NCD1 protomers related by a ~ 2 -fold symmetric dimer interface (Fig. 3a). The A-strand of each protomer protrudes away from the body of that protomer, and into the body of the dimer-related protomer. Because of this exchange of strands, we refer to this interface as the 'strand dimer'. Strand-A kinks away from the G strand, to which it is hydrogen-bonded from residue 7 to 11, at Pro 5–Pro 6 and hydrogen-bonds to strand B' (primes referring to the partner molecule). Thus the β -sheets extend across the interface. The main feature of the strand-dimer interaction is the intercalation of the entire side chain of Trp 2 into the core of the partner molecule in a hydrophobic pocket created by the side chains of Ile 92', Ile 24', Ala 80', Ala 78', Tyr 36', and the aliphatic part of Glu 89'. Trp 2N ϵ of each molecule is hydrogen-bonded to the main-chain carbonyl at position 90 of the partner molecule. The carboxylate of Glu 89 is in position to make a salt bridge to the amino terminus of the native protein, which is artificially extended in our construct. A water molecule (as well as its quasisymmetric mate) is found as a bridging ligand in the strand dimer, tetrahedrally coordinated between Val 30, Arg 23'N, Arg 25'N, and another water molecule. In total, about 1,800 $\text{\AA}^2$ of surface area is buried in the strand-dimer interface.

The protomers of the strand dimer are oriented in parallel, with the corresponding termini of partner molecules pointing in the same directions, as if they were emanating from the same cell surface. This observation and the apparent strength of the interactions of the strand dimer interface have led us to conclude that cadherins probably protrude from the cell surface as dimers. The conserved phenylalanine of the Pro-X-Phe sequence (phenylalanine at positions 108, 223 and 338 in murine N-cadherin), which appears at the end of cadherin domains in most published sequence alignments (for example, see ref. 5), may in fact be at the start of domains in roles analogous to Trp 2. This would require an insertion of two residues in the A-strands of these domains D2–D4 relative to the D1 domain. If this is the case, cadherins may emanate from the cell surface as dimers linked by similar hydrophobic interactions at each domain.

We attempted to characterize the multimerization state of NCD1 by dynamic light scattering and gel-filtration chromatography. In neither case did the estimated relative molecular mass (M_r) depend on Ca^{2+} concentration. Superose-12 chromatography gave an $M_r \sim 21\text{K}$ based on comparative elution times with a set of protein standards, whereas dynamic light-scattering measurements at 0.2–1.0 mM yielded an estimated M_r of 18K–24K. These results suggest that NCD1, which has a molecular mass of 12K, exists as a dimer in solution. The intimate hydrophobic nature of the strand-dimer interface suggests that it is this interface that joins NCD1 monomers in aqueous solution.

The sequence of the part of the A-strand that is exchanged in the strand-dimer interface is identical in N- and E-cadherins, and all contacting groups are identically conserved apart from an Ile 92 $\rightarrow$ methionine change. Two recent studies of recombinantly produced fragments of E-cadherin report that these molecules are monomeric in solution^{19,20}, which is hard to reconcile with our results. It may be that strand dimers of E-cadherin can form only at higher concentrations, or that a cooperative mechanism may be necessary for stability (see later). Nonetheless, we are convinced that the strand dimer will prove to be a universal feature of cadherins.

Adhesion dimer interface

A second large dimer interface with ~ 2 -fold symmetry is also in common among the three lattices (Fig. 3b, c). This interface includes residues from the C, D, F and G strands, as well as the quasi- β helix region of each unit. Protomers of this dimer have an antiparallel orientation, such that the corresponding termini of the partner molecules point in opposite directions, as if they were emanating from opposing cell surfaces. We believe that this interface corresponds to the adhesive interface formed between opposing cells and have therefore named it the 'adhesion dimer'. Specificity-conferring interactions are described later.

The primary areas of contact in the adhesion dimer are located towards the amino terminus of the NCD1 protomers, far from the Ca^{2+} -binding site near the carboxy terminus. This implies that the Ca^{2+} -dependence of cadherin adhesion arises from a mechanism other than from direct regulation of the adhesive contact (see below). The His-Ala-Val sequence in strand F, which is conserved in many D1 domains and has been implicated in cadherin subtype recognition¹⁰, is indeed found on the adhesive face of the molecule. Although His 79 and Val 81 partake in specific interactions across the adhesive interface (see below), the side chain of Ala 80 is buried within the hydrophobic core and does not make contact with the partner molecule. It should be emphasized that the adhesive interface is large and complex; cadherin adhesion cannot be understood by considering the contributions of only a few residues. Also, although there is a superficial resemblance to interfaces in lattice contacts of CD2 crystals²¹, these similarities do not bear up to scrutiny⁵⁰.

Ribbon of adhesive domains

The strand and adhesion dimers, in combination, form a continuous ribbon of protomers. Despite large differences in crystal packing, this ribbon is essentially the same in all three crystal forms of NCD1 (Fig. 3d). The largest difference is that between the type-A and type-C lattices in which there is an r.m.s. deviation of 2.26 $\text{\AA}$ across all four subunits of the translational repeat. The ribbon is composed of NCD1 protomers related to one another by non-intersecting orthogonal diads that correspond to the strand-dimer and adhesion-dimer interactions. Each individual protomer of a strand dimer is joined by an adhesion interface to an opposing strand dimer. This unit repeats to form linear structures of arbitrary length. The carboxy termini of successive strand dimers in a ribbon project out alternately in opposite directions, as if directed towards their respective cell surfaces. We believe this ribbon to be the fundamental unit of cadherin adhesive function.

Ca^{2+} sites between sequential domains

The crystal structures of NCD1 that we have solved have either a UO_2^{2+} or a Yb^{3+} ion bound to a common site near the carboxy terminus of the molecule. In the type-A and B structures, a Yb^{3+} ion is ligated by the side-chain oxygen atoms of Glu 11 and Glu 69, each of which binds the Yb^{3+} in a bidentate fashion; two axial water ligands are also visible (Fig. 4a). The highly incomplete coordination by protein ligands led to high atomic mobility for this Yb^{3+} ion. In the type-C structure, two NCD1 binding sites converge to share in the coordination of a single UO_2^{2+} ion. Asp 67O δ 1 and Glu 11O ϵ 1 from each quasisymmetry

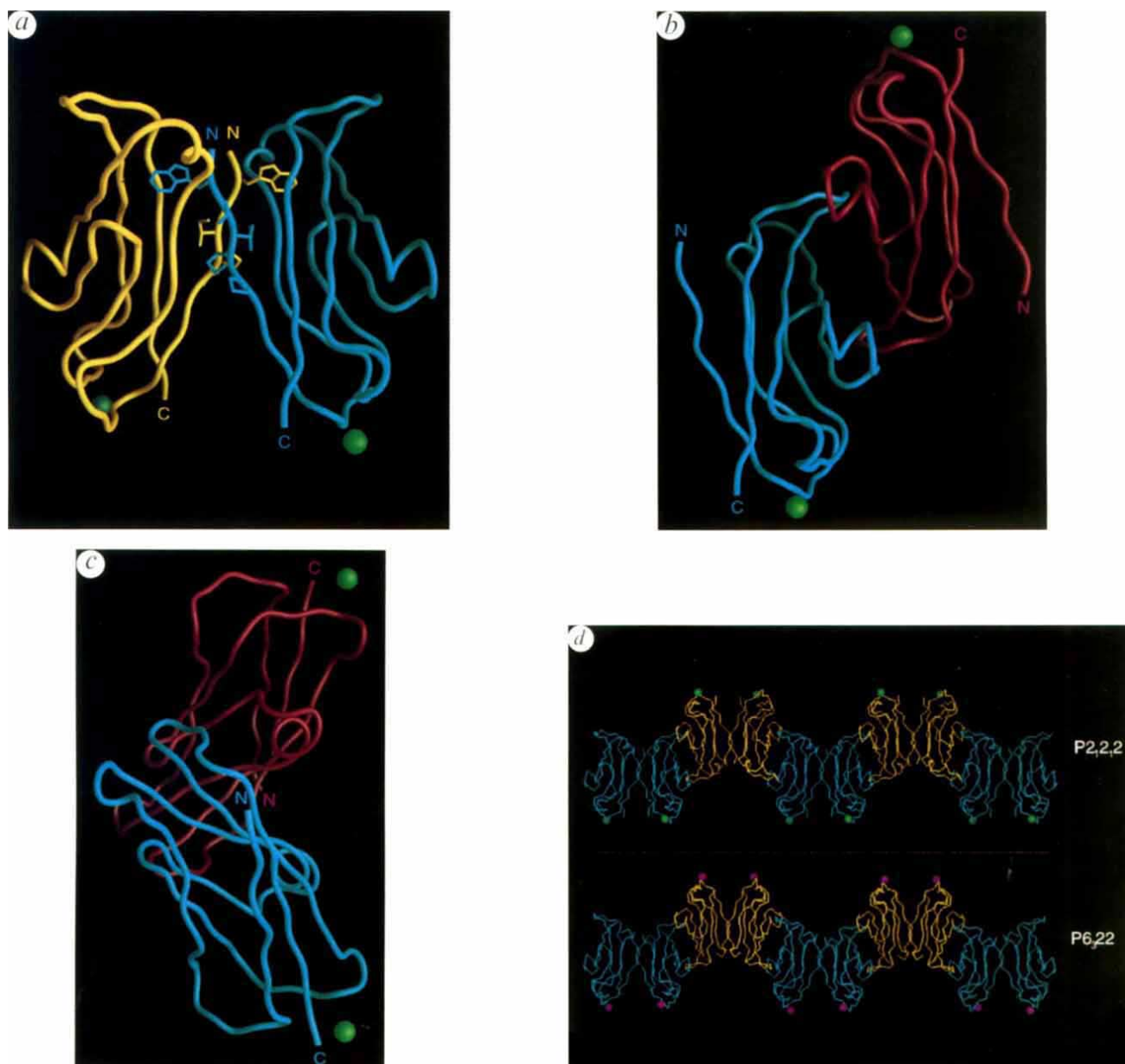


FIG. 3 Molecular associations observed in crystals of NCD1. *a*, Strand dimer. $C\alpha$ worms are shown in blue for molecule 1 and yellow for molecule 2 (type C lattice). The pseudo-two-fold axis runs vertically. Side chains are drawn for residues 2–5 of each protomer. The N and C termini are indicated; uranium atoms are shown as green spheres. Hydrogen bonds that extend the β -character of the A and B' strands are found between Asp 10 and Asp 27'N, and between Val 3N and Arg 25'O (the reciprocal pairs). *b*, Adhesion dimer. Molecule 1 is drawn in blue and a symmetry mate ($x+1/2$, $3/2-y$, $-1-z$) of molecule 2 is drawn in magenta. The quasi-diad axis runs perpendicular to the page. *c*, Adhesion dimer rotated 90° from the view shown in *b*. *d*, Ribbon of adhesive domains. The zipper-like ribbon structure formed by the combination of the adhesion and strand dimer interactions is virtually

mate serve as ligands, together with a single water molecule (Fig. 4b).

The sparse coordination of the metal ions by a single protomer in either crystal form, in conjunction with the position of the metal ion at the carboxy-terminal end of the molecule at the presumptive interface to the next cadherin domain, leads us to speculate that, in intact cadherin molecules, residues from a following domain might complete the sphere of Ca^{2+} coordination. Patterns of sequence conservation superimposed on the NCD1

the same in the UO_2 -complex $P2_12_12$, and the Yb-complex $P6_322$ and $P321$ (not shown) crystal forms. Protomers analogous to cadherins protruding from one cell surface are coloured blue; those protruding from the opposing cell surface are yellow (Fig. 4c). In the Yb-complex crystals, the zipper forms the crystal contacts along the c -axis, which has unit dimension 74.9 Å. For the $P6_322$ crystal, both strand and adhesion dimer interfaces have crystallographic symmetry, and in the $P321$ crystal form only the adhesion dimer is non-crystallographic. In the UO_2 -complex $P2_12_12$ crystals, both interfaces are non-crystallographic and the zipper lies along the a -axis, which has unit dimension 75.2 Å. Ytterbium ions are shown as magenta spheres, and uranium atoms are shown in green; *a*–*c* were made with GRASP⁴⁹; *d* was made with INSIGHT (Biosym Technologies).

structure (Fig. 4c) suggested to us that the native coordination of Ca^{2+} may involve ligands corresponding to Glu 11, Asp 67 and Glu 69 of domain N (which we know donates ligands to the metal ion in NCD1) and the Asp-X-Asp sequence corresponding to residues 27–29 of domain N + 1. This very segment in D2 of E-cadherin has been implicated in Ca^{2+} binding by synthetic-peptide an mutagenesis studies²². Thus Ca^{2+} might function as a structural element in connecting successive cadherin domains. The demonstration that the extracellular

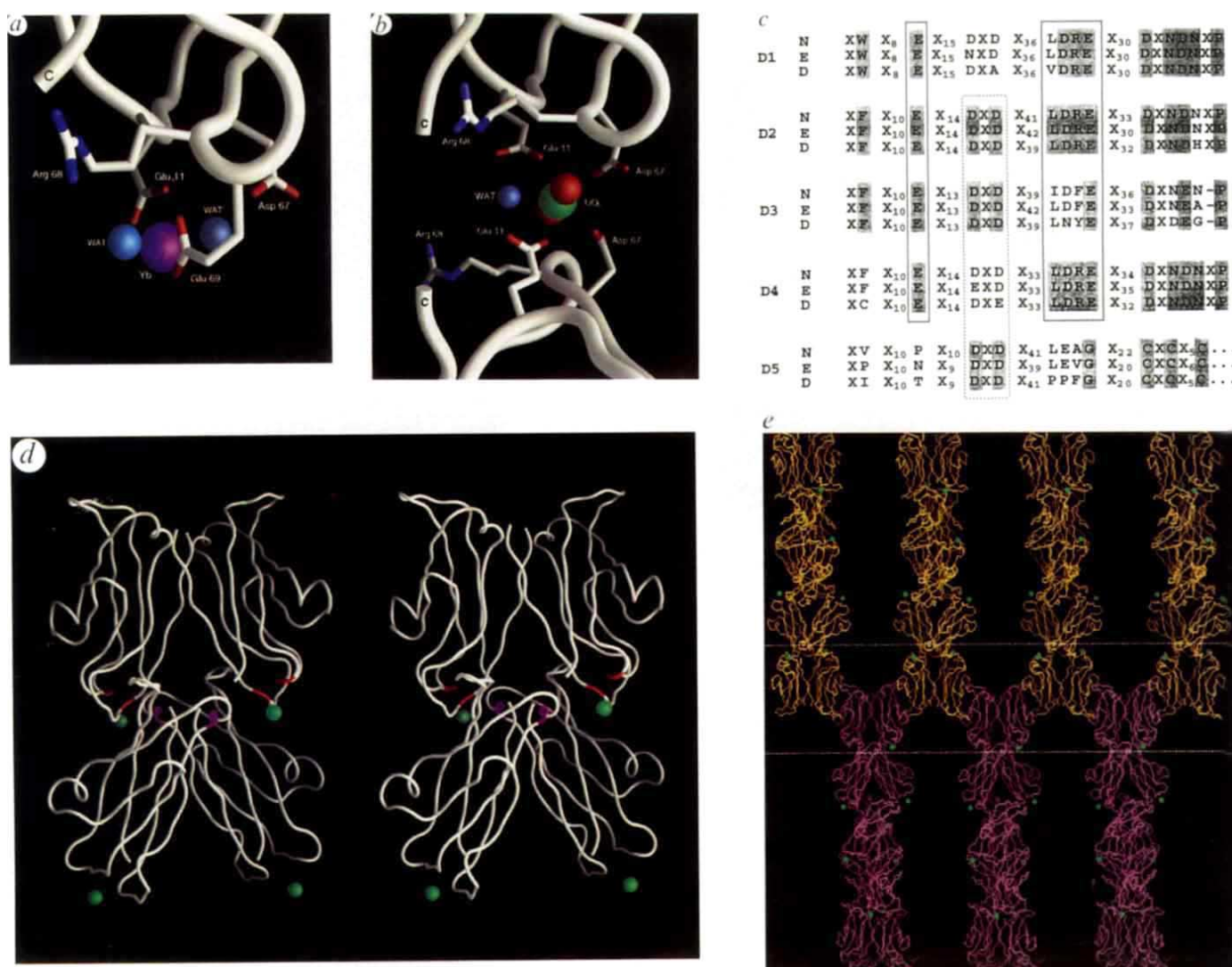


FIG. 4 Ion coordination at the NCD1 Ca^{2+} -binding site. **a**, Yb-complex structure (type A). Ligands to the Yb^{3+} ion, shown as a magenta sphere, are donated by Glu 11, Glu 69 and two visible water molecules, drawn as blue spheres. Asp 67 hydrogen bonds to one of these water molecules. Bonds to oxygen atoms are drawn in red, and nitrogen atoms are in blue. The C terminus of the molecule is indicated. **b**, UO_2 -complex structure. The protomers of molecule 1 and a symmetry mate ($1/2 - x, y - 1/2, 2 - z$) of molecule 2 jointly coordinate a single UO_2 ion, which is drawn in a green and red CPK representation. The uranium atom is coordinated as a pentagonal bipyramid with five equatorial ligands donated by Asp 67 and Glu 11 of molecule 1 and Asp 67' and Glu 11' of molecule 2, and a single water ligand; the covalent oxygen atoms are in axial positions. When the UO_2 - and Yb-complex structures are superposed, the uranium and ytterbium atoms are placed 2.04 Å apart. **c**, Schematic sequence alignment of murine N-cadherin (N) with human E-cadherin (E) and human desmocollin-2 (D). The one-letter amino-acid code is shown for segments of functional residues, including analogues of Trp 2 and possible calcium ligands, and intervening residues are designated by X. Residues that are identical in all three proteins are highlighted. The residues in the solid boxes, which correspond to E11 and LDRE 66–69, denote the metal ligands of NCD1; these are conserved only in D1–D4, and not D5. The DXD motif (also EXD and DXE in D4 domains) is conserved only in D2–D5, and not D1. The DXNDN

linker region between cadherin domains is replaced in the membrane-proximal D5 domain by a cysteine-rich sequence. **d**, Two-domain cadherin model. The α -carbon worm of the strand dimer corresponding to D1 is coloured red at residues that are seen in the NCD1 crystal structures to coordinate a Ca^{2+} analogue (either Yb^{3+} or UO_2^{2+}). The second domain strand dimer unit has been positioned such that the residues corresponding to the conserved DXD sequence of non-D1 cadherin domains (purple) can complete the coordination of the Ca^{2+} ions (drawn as green spheres at the observed Yb positions). **e**, Model of a complete cell-adhesion zipper. NCD1 molecules arranged in the adhesive ribbon structures observed in the crystals were extended to comprise complete five-domain cadherin extracellular segments as described in the text. Molecules in yellow correspond to complete cadherin extracellular segments emanating from one cell surface. Molecules drawn in purple correspond to cadherin extracellular segments emanating from the opposing cell surface. The membrane-to-membrane distance predicted by this model is ~ 290 Å. The distance between 2-fold axes of adjacent cadherin dimers protruding from the same cell surface is ~ 75 Å. The structure within the dotted-line box is the experimentally observed ribbon from the crystals of NCD1. The structure outside the box is modelled. Panels **a–d** were made with GRASP²⁹; **e** was made with INSIGHT.

domain of E-cadherin is a stiff rod in the presence of Ca^{2+} which collapses in its absence¹⁹, supports a role for Ca^{2+} in stabilizing interactions between successive cadherin domains.

We have constructed multidomain models based on the conjecture that each domain will be involved in a strand-dimer interaction, and with the simplifying assumption that the 2-fold

strand-dimer axes of all domains will coincide. Strand-dimer units were stacked head to tail (N terminus to C terminus; 34.0 Å translation), and rotated (120° about the strand-dimer axis) so that the aspartate residues at positions 27 and 29 were near the metal ions bound to the preceding strand dimer unit (Fig. 4d). In reality, we expect some conformational change upon calcium

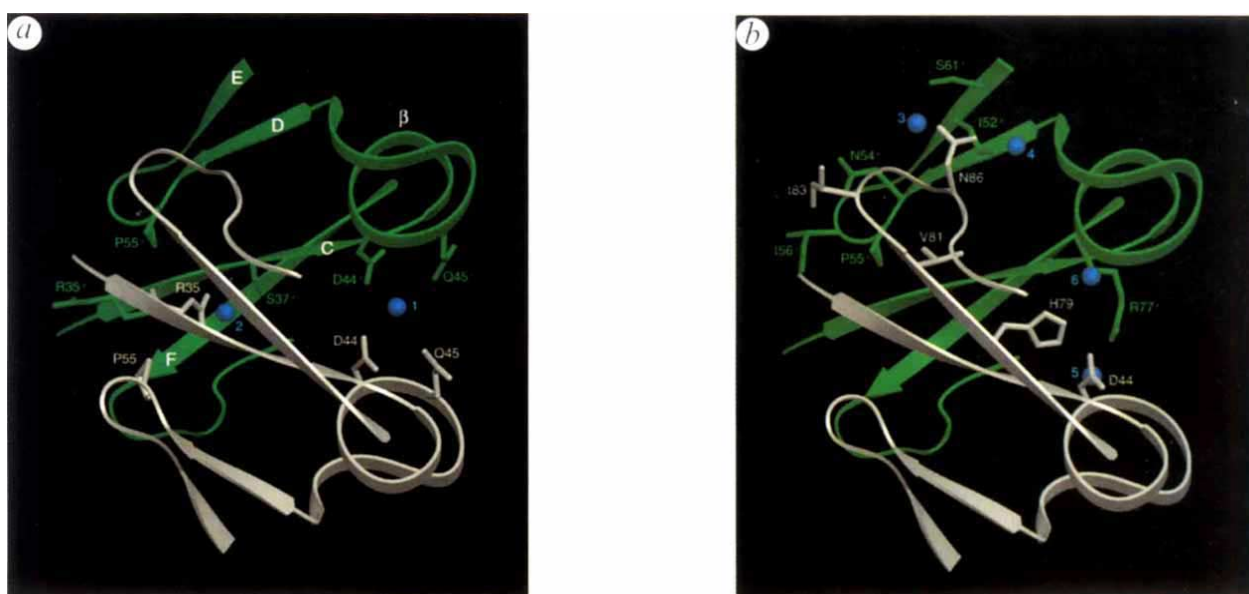
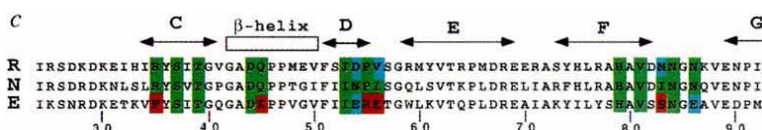


FIG. 5 Adhesive interface interactions. Ribbon diagrams in *a* and *b* encompass molecular portions involved in the adhesive interface. Water molecules are depicted as blue spheres. Type C molecule 1 is shown as grey and its partner, a symmetry-mate ($x+1/2$, $3/2-y$, $-1-z$) of molecule 2, is drawn in green. *a*, Side chains for residues involved in quasi- β helix interactions and C strand to C strand interactions are shown. The quasi- β helix regions of the partner molecules interact sparsely through a single water-mediated contact. Asp 44O δ 2 and Gln 45N ϵ 2 from molecule 1 and Asp 44'O δ 2 and Gln 45'N ϵ 2 from molecule 2 tetrahedrally coordinate a common water molecule (labelled 1). Asp 45 is conserved in D1 domains, and position 46 is generally occupied by a side chain that can participate in hydrogen bonding. Most interactions in the C strand centre around the side chain of Arg 35. This is the most notable site of asymmetry in the adhesion dimer. The Arg 35 side chain points downwards (towards the carboxy terminus) and into the centre of the interface where N η 2 hydrogen bonds to Tyr 36'O, whereas the Arg 35' side chain points upwards and out towards solvent. The aliphatic stalks of Arg 35 and 35' pass by each other in a hydrophobic contact, and the β - and γ -carbon atoms of Arg 35 and 35' make contacts to the proline rings of 55' and 55 respectively. Arg 35 N η 2 jointly coordinates a water molecule (labelled 2) with the side-chain hydroxyl of Ser 37'. *b*, Side chains for residues involved in DE-loop to FG-loop interactions are shown. This is the most extensive area of contact in the adhesion dimer and, for clarity, side chains involved in near-symmetric interactions are shown for only one molecule. Hydrophobic contacts are observed between the side chains of Ile 83 and Ile 56' at the apex of the DE loop, between Val 81 and Pro 55', and between the α -carbon of Gly 85, which is conserved in all D1 domains, and the Pro 55' ring. A direct hydrogen bond between Asn 54'N δ 2 and Ile 830 is also observed. An interesting van der Waals contact is seen between



the polar Asn side chain at 86 and the apolar Ile side chain at 52', with Asn 86O δ 1 3.4 Å from C δ 1 and 3.9 Å from C γ 2 of Ile 52'. Another contact of this type is observed between carbonyl Ile 53'O and Gly 85 C α , which are 3.3 Å apart. Direct interactions between these loops are largely symmetric in the partner molecules. There are several water-mediated interactions and many of these are asymmetric. One water molecule (labelled 3) is tetrahedrally coordinated between Gln 59'O, 840 Ser 61'O γ , and another water molecule. The reciprocal interaction is distinctive and is mediated through two water molecules (not shown). The single-water-mediated interaction (labelled 4) between Asn 86O δ 1 and Gly 49'O is completely absent in the quasi-symmetric case. Another water molecule (labelled 5) bridges Asp 440 and Arg 77'N η 1. The reciprocal interaction is not observed, but rather His 79N ϵ 2 hydrogen-bonds to a water (labelled 6) bound to Asp 44'O. There are no true salt bridges observed in the adhesion interface, but some long-range favourable electrostatic contacts are observed. Arg 77'N η 2 comes within 3.6 Å of Asp 44 O δ 2, but the reciprocal interaction is not observed because the side chain of Arg 77 points down and out of the interface. His 79N ϵ 2 is 4.5 Å from Asp 44'O δ 2, and the reciprocal distance is 4.3 Å. Figures made with SETOR⁴⁷. *c*, Comparison of sequences involved in forming the adhesive interface in N-, R- and E-cadherin. Residues whose side chains are directly involved in adhesive contacts are highlighted. Those residues that are identical to the corresponding residue in N-cadherin are shown in green. Conservative substitutions (hydrophobic to hydrophobic for example) are shown in blue, and non-conservative substitutions in red.

binding. The Leu-Asp-Arg-Glu (LDRE)-containing loop differs in type-A and type-C lattices, and in the BC loop, which is one of the most highly disordered regions of the molecule.

The linker region between cadherin domains, which is disordered in our one-domain structure, contains a conserved Asp-X-Asn-Asp-Asn motif. The abundance of possible ligands in this sequence suggests a further involvement in Ca²⁺ binding. One intriguing possibility, suggested by our multidomain model, envisages that the two linker regions of strand-dimer-associated domains pass by each other to coordinate jointly a single Ca²⁺ ion situated on the 2-fold axis of the strand dimer. The Asp-X-Asn-Asp-Asn motif is not found in the corresponding region at the base of the membrane-proximal D5 domain. Rather, a pattern of three cysteine residues, Cys-X-Cys-X₂-Cys, is conserved in many cadherins. The odd number of cysteines in each D5 domain precludes the possibility that they can all be involved in

intradomain disulphide bonds. We therefore suggest that some of these cysteines may covalently link the two D5 domains across the strand-dimer axis at the base of the cadherin molecule.

Cell-adhesion zipper

The combination of the ribbon of adhesive domains seen in these crystal structures together with our model of a Ca²⁺-rigidified succession of domains for dimeric cadherin molecules generates an interdigitated entity (Fig. 4e) which we call the 'cell-adhesion zipper'. The membrane-to-membrane distance measured directly from this model is ~290 Å, which agrees well with the range of estimates of intermembrane separations of ~200–350 Å reported from electron microscopy of desmosomes and adherens junctions^{23, 27}. This observation alone, given our knowledge of the overall structure of cadherin domains, implies that there are terminal contacts between extended molecules. Note also that

the essence of the cell-adhesion zipper is preserved even if our hypothetical model for Ca^{2+} -mediated domain association is wrong. Whatever the connection to the membrane is like, it is the teeth of the zipper seen in the ribbon that provide the mechanism for cell adhesion.

Our solution experiments show that NCD1 remains a dimer at concentrations up to 1 mM. If we assume that the solution dimer is the strand dimer, then the adhesive interface must be intrinsically weak. Despite the large surface area, most interactions are long-range or water-mediated in character. Although the adhesive interface buries about 3,300 Å² surface area in total, one third of this area is lost if water molecules are removed. Moreover, the interface is intrinsically asymmetric, with several quasi-diad-related side chains found in widely differing conformations. The adhesion zipper provides a mechanism by which such weak adhesive interactions can be compounded to form strong intercellular bonds.

The linear nature of the zipper structure fits well with our understanding of cadherin function. Cadherins form the intercellular bonds in both *zonula adherens* and desmosome junctions. The classic cadherins are known to associate indirectly through their cytoplasmic domains with the actin filaments of *zonula adherens*^{13,14,28}. In epithelia these structures comprise belts^{23,29} in which linear actin filaments and associated cadherins encircle the cells. The linear zipper structure may associate with the linear cytoplasmic filament structure of the adhesion belt. Likewise, cytoplasmic tonofilaments in the desmosomal structure run through the desmosome parallel to the surface of the membrane^{30,32}. Linear zippers of desmosomal cadherins may associate along the length of these filaments as well.

This raises the question of whether the cytoplasmic filament system mediates cell adhesion by positioning extracellular cadherin zipper structures, or whether it is the zipper structures that guide the intracellular structures. Evidence exists that supports both hypotheses, and we think both mechanisms are operative. On the one hand, the cadherin cytoplasmic domain and its interaction with other proteins is necessary for cadherins to function in cell adhesion^{15,16,33}; on the other, disruption of the extracellular function of cadherins, either by Ca^{2+} deprivation^{34,35} or by treatment with anti-cadherin antibodies³⁴, can lead to the disintegration of the microfilament bundles of the adhesion belt. We suggest that these observations result from an intimate coupling of the extracellular zipper structure with a superstructure of cytoplasmic components of *zonula adherens*. We expect that this is the case for desmosomal structures as well.

The structure of the cell-adhesion zipper has a much greater density of protein at the ribbon of adhesive domains in the centre of the intercellular space than it has elsewhere. Features consistent with this are seen in electron micrographs, namely the 'central dense stratum' commonly observed in desmosomes^{23,32} and less commonly in *zonula adherens*³⁶. The readier visualization of this medial density in desmosomes may indicate that the concentration of cadherins is greater in these junctions.

Adhesive specificity

Interactions in the adhesion dimer can be grouped into three primary areas: (1) β -helix-to- β -helix contact; (2) C-strand-to-C-strand contact; and (3) FG-loop-to-DE-loop contact (Fig. 5 legend). Sequence alignment of classic cadherins, desmocollins and desmogleins demonstrates that there are usually zero, or at most two, insertions or deletions of amino acids in D1 domains relative to NCD1. Therefore, the structure of NCD1 provides a picture of a scaffolding that is probably common to many cadherin adhesive interactions, and defines those residues that determine cadherin adhesive specificity.

Comparison of the sequences of N- and R-cadherins, which (unlike most cadherins) interact heterophilically³⁷, shows that only three changes, all conservative, occur in residues involved in the adhesive interface (Fig. 5c). These are Asn 54→Asp,

Ile 56→Val and Ile 83→Met, which might support favourable heterophilic interactions. Analysis of the sequence of E-cadherin, which does not interact with N-cadherin, shows that it has a more significant alteration of the residues in the adhesive interface. These include three conservative and four non-conserved substitutions. To maintain self-adhesiveness, complementary pairs of non-conservative side-chain substitutions would be predicted, as seems to be the case. For example, both residues involved in the hydrophobic contact between Ile 56 and Ile 83 in N-cadherin are replaced in E-cadherin by polar residues glutamate and serine, respectively, which could hydrogen-bond. Another non-conservative substitution, Arg 35→Phe, is itself a 'pair' because Arg 35 interacts with Arg 35' across the quasi-2-fold of the adhesion dimer in a hydrophobic interaction. It appears that the overall features of the cadherin adhesive interface will be similar among many members of the cadherin family. In particular, we predict from sequence comparison that this will also be the mode of adhesion by desmosomal cadherins. Despite the large area of the adhesive interface, the basis for specificity of cadherin interactions may be subtle; changes at only a few residues may alter the balance in these intrinsically weak interfaces.

To our knowledge, the structure of NCD1 provides the first atomic-level view of a biologically specific cell-adhesion interaction. Crystals in three different lattices each recapitulate the interactions of cadherins functioning at the interface between adherent cells. The crystal structures suggest that cadherin-mediated cell adhesion is not based solely on the stability of association between individual molecules, but rather that opposing cell surfaces are joined in a multimeric superstructure that provides a cooperative mechanism to form strong intercellular bonds. Many cell-adhesion molecules from outside the cadherin superfamily can associate only weakly as monomers³⁸. Cooperative superstructures, like the cadherin cell-adhesion zipper, may also be a feature in other adhesive interactions between cells. Finally, an NMR study of the N-terminal domain D1 from murine E-cadherin has now been reported⁵¹. □

Received 13 January; accepted 17 February 1995.

- Wilson, H. V. *J. exp. Zool.* **5**, 245–258 (1907).
- Galtsoff, P. S. *J. exp. Zool.* **42**, 183–222 (1925).
- Oda, H., Uemura, T., Harada, Y., Iwai, Y. & Takeichi, M. *Dev. Biol.* **165**, 716–726 (1994).
- Takeichi, M. *Science* **251**, 1451–1455 (1991).
- Hatta, K., Nose, A., Nagafuchi, A. & Takeichi, M. *J. Cell Biol.* **106**, 873–881 (1988).
- Nose, A., Tsuji, K. & Takeichi, M. *Cell* **61**, 147–155 (1990).
- Takeichi, M., Inuzuka, H., Shimamura, K., Fujimori, T. & Nagafuchi, A. *Cold Spring Harbor Symp. quant. Biol.* **LV**, 319–325 (1990).
- Takeichi, M. *A. Rev. Biochem.* **59**, 237–252 (1990).
- Amagai, M., Karpatti, S., Prussick, R., Klaus-Kovtun, V. & Stanley, J. R. *J. clin. Invest.* **90**, 919–926 (1992).
- Blaschuk, O. W., Sullivan, R., David, S. & Pouliot, Y. *Dev. Biol.* **139**, 227–229 (1990).
- Takeichi, M. *J. Cell Biol.* **75**, 464–474 (1977).
- Yoshida, C. & Takeichi, M. *Cell* **28**, 217–224 (1982).
- Magee, A. I. & Buxton, R. S. *Curr. Opin. Cell Biol.* **3**, 854–861 (1991).
- Cowin, P. *Proc. natn. Acad. Sci. U.S.A.* **91**, 10759–10761 (1994).
- Kintner, C. *Cell* **69**, 225–236 (1992).
- Fujimori, T. & Takeichi, M. *Molec. Biol. Cell* **4**, 37–47 (1993).
- Laclaire, J. L. *J. Synchr. Rad.* **1**, 12–18 (1994).
- Hendrickson, W. A. *Science* **254**, 51–58 (1991).
- Pokutta, S., Herrenknecht, K., Kemler, R. & Engel, J. *Eur. J. Biochem.* (1994).
- Tong, K. I. et al. *FEBS Lett.* **352**, 318–322 (1994).
- Bodian, D. L., Jones, E. Y., Harlos, K., Stuart, D. I. & Davis, S. J. *Structure* **2**, 755–766 (1994).
- Ozawa, M., Engel, J. & Kemler, R. *Cell* **63**, 1033–1038 (1990).
- Farquhar, M. G. & Palade, G. E. *J. Cell Biol.* **17**, 375–412 (1963).
- Overton, J. *Prog. Surf. membr. Sci.* **8**, 161–208 (1974).
- McNutt, N. S. & Weinstein, R. S. *Progr. Biophys. molec. Biol.* **26**, 45–101 (1973).
- Campbell, R. D. & Campbell, J. H. *Origin and Continuity of Desmosomes* (Springer, New York, Heidelberg and Berlin, 1971).
- Staehelein, L. A. *Int. Rev. Cytol.* **39**, 191–283 (1974).
- Nagafuchi, A., Tsukita, S. & Takeichi, M. *Sem. Cell Biol.* **4**, 175–181 (1993).
- Wessels, N. K. et al. *Science* **251**, 135–143 (1991).
- Fawcett, D. W. *An Atlas of Fine Structure* 365–382 (Saunders, New York, 1958).
- Rogers, D. C. *J. Ultrastruct. Res.* **13**, 478–499 (1965).
- Kelly, D. E. *J. Cell Biol.* **28**, 51–72 (1966).
- Matsuyoshi, N. et al. *J. Cell Biol.* **118**, 703–714 (1992).
- Volberg, T., Geiger, B., Kartenbeck, J. & Franke, W. J. *Cell Biol.* **102**, 1832–1842 (1986).
- Voit, T. & Geiger, B. *J. Cell Biol.* **103**, 1451–1464 (1986).
- McNutt, N. S. *Amer. J. Cardiol.* **25**, 169 (1970).
- Inuzuka, H., Miyatani, S. & Takeichi, M. *Neuron* **7**, 69–79 (1991).
- van der Merwe, P. W. & Barclay, A. N. *Trends biochem. Sci.* **19**, 354–358 (1994).

39. Smith, D. B. & Johnson, K. S. *Gene* **67**, 31–40 (1988).
 40. Gröbel, G., Als-Nielsen, J. & Freund, A. K. *J. De Physique IV, Colloque C9* **4**, 27–34 (1994).
 41. Freund, A. K. *SPIE Proc.* **1997**, 223 (1993).
 42. Otwinowski, Z. in *Data Collection and Processing* (eds Sawyer, L., Isaacs, N. & Bailey, S.) 56–62 (SERC Daresbury Laboratory, Warrington, UK, 1993).
 43. CCP4: A suite of Programs for Protein Crystallography (SERC Collaborative Computing Project no. 4, Daresbury Laboratory, Warrington, UK, 1979).
 44. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. *Acta crystallogr. A* **47**, 110–119 (1991).
 45. Brünger, A. T. *X-PLOR, Version 3.1, A System for X-ray Crystallography and NMR* (Yale Univ. Press, New Haven, USA, 1992).
 46. Kabsch, W. & Sander, C. *Biopolymers* **22**, 2577–2637 (1979).
 47. Evans, S. V. *J. molec. Graph.* **11**, 134–138 (1993).
 48. Kraulis, P. J. *J. appl. Crystallogr.* **24**, 946–950 (1991).
 49. Nicholls, A., Sharp, K. A. & Honig, B. *Proteins* **11**, 281–296 (1989).
 50. Shapiro, L., Kwong, P. D., Fannon, A. M., Colman, D. R. & Hendrickson, W. A. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
 51. Overduin, M. et al. *Science* **267**, 386–389 (1995).

ACKNOWLEDGEMENTS. We thank J. Laclaire and the accelerator group at ESRF for their efforts with the undulator, members of the EMBL at Grenoble for loan of their MAR detector, and C. Ogata, X. Zhao and Y. Liu for their help with data collection at Brookhaven. L.S. was supported in part by a training grant from the National Eye Institute of the NIH. A.M.F. received fellowships from the Human Frontiers Science Program Organization and the National Multiple Sclerosis Society. This research was supported in part by grants from the NIH to W.A.H. and D.R.C. and from the NSF to W.A.H. Beamline X4A at the National Synchrotron Light Source, a DOE facility, is supported by the Howard Hughes Medical Institute. Coordinates have been deposited in the Brookhaven Protein Data Bank.

LETTERS TO NATURE

Nucleosynthesis of ^{11}B -rich boron in the pre-solar cloud recorded in meteoritic chondrules

Marc Chaussidon* & François Robert†

* CRPG-CNRS, BP 20, 54501 Vandoeuvre-lès-Nancy, France

† Muséum National d'Histoire Naturelle, 61 rue Buffon, 75015 Paris, France

MODELS of the chemical evolution of the Galaxy, in which most elements are created inside stars and distributed by stellar winds and supernovae, cannot produce the observed abundances of boron and beryllium¹. These elements have been produced continuously since the Big Bang by collisions between Galactic cosmic rays (very energetic protons and α -particles) and heavier elements, such as carbon and oxygen, in the interstellar medium^{2–6}. But models of chemical evolution that include these effects predict a boron isotope ratio ($^{11}\text{B}/^{10}\text{B} = 2.5$, ref. 2) that is very different from that observed on Earth and in meteorites ($^{11}\text{B}/^{10}\text{B} \approx 4.0$, refs 7–9). Here we present ion-probe measurements of the $^{11}\text{B}/^{10}\text{B}$ ratio in meteoritic chondrules, which reveal significant variations (3.84–4.25) correlated with the beryllium and boron concentrations. These correlations can be explained by production of ^{11}B -rich boron in the pre-solar cloud, resulting from collisions between interstellar hydrogen (and helium) and low-energy cosmic rays¹⁰ such as the carbon and oxygen nuclei recently observed in the Orion star-forming complex¹¹. Our results also suggest that isotopic heterogeneities have been partially preserved during the process of chondrule formation.

This study was done on four different chondrites (Allende CV3, Orgueil CI, Semarkona LL3, Hedjaz L3–5) which were selected to represent a wide range of compositions and conditions of formation. Orgueil represents the least chemically fractionated sample relative to the Sun¹² and Allende is well known to be rich in supernovae condensates^{13,14}. Semarkona has the highest pristine D/H ratios of the Solar System¹⁵; such a

deuterium enrichment is interpreted as the result of the chemistry taking place in the interstellar cloud that gave birth to the solar nebula¹⁶ and testifies to the occurrence of well preserved pre-solar material. Individual chondrules were analysed because chondritic meteorites are made of up to 85% of chondrules which were partially molten during flash heating events occurring in the nebula during the very first stages of Solar System formation¹⁷. Although the low-temperature matrix has been subjected to weathering during secondary parent body processes¹⁸, high-temperature chondrules contain relict grains out of chemical equilibrium with other minerals and therefore may have recorded the chemical heterogeneities of the protosolar nebula¹⁹.

Be and B contents and B isotopic ratios were determined by ion probe (Cameca ims3f at CRPG-CNRS, Nancy) following the procedures developed for low beryllium and boron contents in mantle rocks^{9,20}. Boron isotopes were measured at $\Delta M/M = 1,800$ (thus resolving all possible isobaric interferences) and boron contents with an additional energy filtering of $-60 \pm 10\text{V}$ (see refs 9 and 20 for details). This approach allows *in situ* analyses at a scale of $\sim 25\text{ }\mu\text{m}$ and avoids most of the analytical problems encountered in classical mass spectrometric analyses, namely the possible terrestrial contamination²¹ and the required high chemical purification of boron extracts. Boron isotope ratios are reported in $\delta^{11}\text{B}(\text{‰})$ notation ($\delta^{11}\text{B} = 1,000 \times ((R_{\text{sp}}/R_{\text{std}}) - 1)$ where R_{sp} and R_{std} are the $^{11}\text{B}/^{10}\text{B}$ ratios of the sample and of the NBS 951 international standard, $R_{\text{std}} = 4.04558$). Boron and beryllium contents are given as atomic ratios to silica (within $\pm 10\%$), and $\delta^{11}\text{B}$ values are determined within $\pm 5\%$ (for $\text{B}/\text{Si} > 2 \times 10^{-5}$) and within $\pm 10\%$ for lower boron contents ($\pm 1\sigma$).

The data obtained across the three individual chondrules shown in Fig. 1 were selected for their large variations in B/Si and Be/Si atomic ratios and in $\delta^{11}\text{B}$ values. Orgueil, which is chondrule free, is also reported for solar reference value. Chondrules have $\delta^{11}\text{B}$ values which range between $\sim -50\%$ and $\sim +50\%$ and are correlated with the boron contents. These values encompass the range previously reported for bulk chondrites (three $\delta^{11}\text{B}$ values between -57 and -35% (ref. 7) and six between -8.5 and $+1.8\%$, ref. 8) but no systematic chemical or spatial distribution of $\delta^{11}\text{B}$ values has been found. They far exceed the $\delta^{11}\text{B}$ values measured in fresh mantle rocks (from

TABLE 1 Production fluxes of B calculated in the Orion nebula

Incident particle	Target	Product	Cross-section (mb)*	N (atoms s ⁻¹)	N/Si† (s ⁻¹)	N/Si† (s ⁻¹)
^{12}C	H	$^{12}\text{C}^{\dagger}$	200	2.0×10^{39}	4.7×10^{-19}	4.7×10^{-20}
^{12}C	H	^{11}B	100	1.0×10^{39}	2.3×10^{-19}	2.3×10^{-20}
^{12}C	H	^{10}B	22	2.2×10^{38}	5.1×10^{-20}	5.1×10^{-21}
^{16}O	H	^{11}B	10	1.0×10^{38}	2.3×10^{-20}	2.3×10^{-21}
^{16}O	H	^{10}B	2.2	2.2×10^{37}	5.1×10^{-21}	5.1×10^{-21}

* The cross-sections at 10–30 MeV per nucleon are taken from ref. 1 ($1\text{ mb} = 10^{-27}\text{ cm}^2$).

† The B/Si ratio is calculated assuming that 10% (first column) and 100% (second column) of the available cosmic Si is in the gas phase.

‡ ^{12}C atoms in an excited state.

TABLE 2 B contents and isotopic compositions calculated for the Orion nebula compared with observed values for GCR, meteorites and the Earth

Values calculated for the Orion nebula						
Irradiation time (Myr)	0*	0.4	0.6	1.2		
$\delta^{11}\text{B}$ (‰)	-400	-40	0	+49		
B/Si (100% Si in the gas)	1×10^{-6}	4.90×10^{-6}	6.84×10^{-6}	1.27×10^{-5}		
B/Si (10% Si in the gas)	1×10^{-5}	4.80×10^{-5}	6.70×10^{-5}	1.24×10^{-4}		
Values measured for GCR, meteorites and the Earth						
	GCR	Semarkona	Allende	Hedjaz	Orgueil	Earth
$\delta^{11}\text{B}$ (‰)	-400	-30 to +40	-50 to +20	-40 to +40	-2 ± 5	-7 ± 5
B/Si	—	8×10^{-6} to 2×10^{-5}	2×10^{-5} to 7×10^{-5}	1×10^{-4} to 8×10^{-4}	2.6×10^{-5}	3×10^{-6}

* For time 0, B contents and isotopic compositions are those calculated for high-energy GCR². Data for GCR from ref. 33, for the silicate Earth from ref. 9 and for meteorites from this study.

-12 ± 2 to $-2 \pm 2\%$, ref. 9) implying that (1) laboratory contamination is ruled out because terrestrial samples with low boron contents were prepared in a similar way to meteorite samples and that (2) high-temperature reactions between common mafic minerals cannot produce large isotopic fractionations. In fact no terrestrial process has been shown to be able to yield such variations within a single rock. The B/Si ratios range between 3×10^{-6} and 2×10^{-4} (between 0.25 and 16 p.p.m. for 45 wt% SiO_2) with one extreme value at 1×10^{-3} . This range is compatible with the domain defined by literature data²²⁻²⁴, and shows that chondrules can be depleted or enriched relative to Orgueil (2.6×10^{-5} our data; 2.1×10^{-5} from ref. 10), that is, to the Sun (1.1×10^{-5} based on solar photospheric determination¹⁰). As for many other elements, Orgueil represents a 'mean' solar boron, both for its abundance and isotopic composition ($\delta^{11}\text{B} = -2 \pm 5\%$). Be/Si ratios range between 9×10^{-8} and 1×10^{-4} and are correlated with B/Si ratios giving a fairly constant B/Be ratio of 17.2 ± 5.8 (see Fig. 1). The constancy of the B/Be ratio

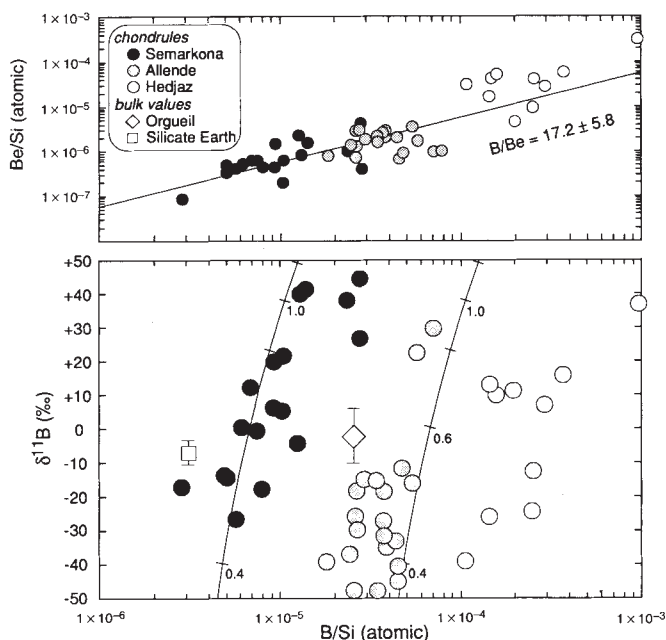


FIG. 1 Ion-probe determinations of the atomic B/Si ratio versus Be/Si and $\delta^{11}\text{B}$ (‰) for excentroradial chondrules from Semakona and Hedjaz meteorites and for an olivine barred chondrule from Allende. The mean $\delta^{11}\text{B}$ value of Orgueil (five determinations) and the value for bulk silicate Earth⁹ are shown for reference. The chondrules have a B/Be ratio (17.2 ± 5.8) close to the theoretical spallation ratio². The large variations in $\delta^{11}\text{B}$ values and their correlations with B/Si, may result from a massive production of ^{11}B -rich boron in the pre-solar interstellar cloud. The relationships between $\delta^{11}\text{B}$ and B/Si, which can be calculated from spallation rates observed in the dense cloud of Orion (see text, Fig. 2 and Table 2), are shown as solid curves with irradiation times indicated as Myr.

demonstrates that these elements were not significantly fractionated during their incorporation into chondrules although beryllium is supposed to be more refractory than boron^{25,26} and chemical fractionation occurs during melt-crystal equilibrium²⁰. The actual occurrence of isotopic variations in chondrules reflects either a large initial isotopic heterogeneity or a lack of isotopic homogenization between unmelted minerals or both.

Isotopic fractionation, often invoked in Solar System processes, is unable to explain $\delta^{11}\text{B}$ variations of $\sim 100\%$ and their positive correlations with B/Si ratios (Fig. 1). Some examples follow. (1) Boron devolatilization experiments of synthetic chondrules fused by a laser beam under vacuum, show isotopic fractionations $\leq 7\%$, even for degassed fractions of $\sim 50\%$ (M.C. and F.R., unpublished results). In addition, by analogy with what has been observed in evaporation experiments for many elements²⁷, ^{11}B would be enriched in solid residues resulting in a negative correlation with B/Si ratios, opposite to that observed (Fig. 1). (2) Rims of flanged tektites (terrestrial impactite) which are remelted and significantly outgassed in the high terrestrial atmosphere show boron isotopic effects $\leq 3\%$ (ref. 28). (3) Volcanic gases exhibit $\delta^{11}\text{B}$ variations $\leq 15\%$ (refs 29, 30). (4) Apart from fractionations associated with planetary processes, the boron isotopic fractionation taking place during ion-molecule reactions in dense interstellar clouds appears, in contrast to hydrogen, to be close to 0‰. This is shown by the fairly constant $\delta^{11}\text{B}$ value (from -5 to $+10\%$; six determinations in the immediate vicinity of the chondrule) measured for the deuterium-rich phyllosilicates whose high D/H ratio probably results from interstellar chemistry (D/H up to 7×10^{-4} , that is, a deuterium enrichment relative to the interstellar hydrogen by a factor of ~ 35)¹⁶. Finally, note that a local boron contribution by the surrounding matrix is also unlikely because the phyllosilicates in Semakona, although boron-rich (B/Si = 1×10^{-3}), exhibit $\delta^{11}\text{B}$ values around $+3\%$ and thus cannot be a source of contamination for chondrules whose $\delta^{11}\text{B}$ values range from -35 to $+45\%$.

Most probably, the large $\delta^{11}\text{B}$ variations found in chondrules are the result of variations in the amount of processing, through spallation reactions, in the molecular cloud from which the Solar System formed. This is suggested by the agreement between the theoretical spallogenic B/Be ratio (ranging from 15 to 22)² and that observed in chondrules (B/Be = 17.2 ± 5.8). Reactions between meteorites and Galactic cosmic rays (GCR) in the present-day Solar System would yield $^{11}\text{B}/\text{Si}$ ratios $\leq 3 \times 10^{-9}$ (with a typical exposure age $\leq 10^7$ years for chondrites and a solar flux of $50 \text{ protons cm}^{-2} \text{ s}^{-1}$), which is 3 orders of magnitude lower than the observed value. Therefore boron isotopic variations must reflect a pre-accretionary isotopic heterogeneity, probably carried by grains with diameters $\sim 5 \mu\text{m}$. This size accounts for the observed $\delta^{11}\text{B}$ variations of $\pm 40\%$ around their mean value, for a $500 \mu\text{m}^2$ spot size and assumes a mixing between two isotopic components having $\delta^{11}\text{B}$ values of $+100$ and -400% (see next paragraph for these $\delta^{11}\text{B}$ estimates).

In the Orion molecular cloud, γ -ray lines at 4.4 and 6.1 MeV have been recently observed and ascribed to the interaction between heavy cosmic rays (^{12}C and ^{16}O nuclei) and interstellar

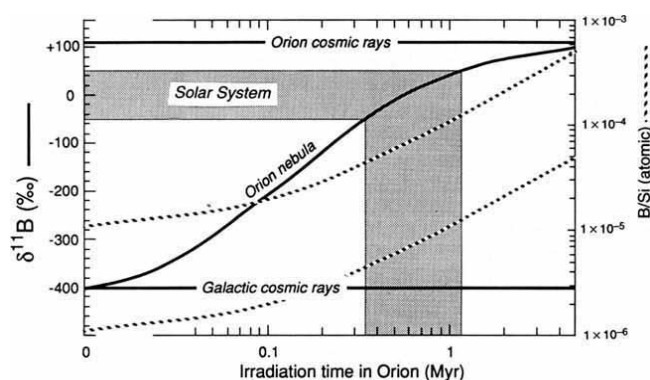


FIG. 2 The relationships between $\delta^{11}\text{B}$ and B/Si with irradiation time are shown as solid and dashed curves, respectively. These variations are calculated from spallation rates in Orion (see text, Fig. 2 and Table 2) involving a low-energy flux (10–30 MeV per nucleon) of heavy cosmic rays (^{16}O and ^{12}C) interacting with interstellar H. The evolution of $\delta^{11}\text{B}$ results from the progressive mixing with time between GCR boron ($\delta^{11}\text{B} \sim -400\text{‰}$ observed³³ and calculated²) and 'planetary boron' ($\delta^{11}\text{B} \sim +110\text{‰}$ at 10–30 MeV, refs 2, 5). Increase with time of the B/Si ratio in the gas is shown for 100% (lower dashed curve) and 10% of the cosmic Si in the gas (upper dashed curve). In Orion, irradiation times between ~ 0.3 and ~ 1.1 Myr are required to produce B/Si and $\delta^{11}\text{B}$ similar to those observed in meteorites.

H (refs 11, 31). Clayton³² showed that several short-lived isotopes (such as ^{26}Al or ^{129}I) can be produced in this environment by spallation reactions. In this example, it is of particular interest to note that the interaction between H and low-energy cosmic rays (10–30 MeV per nucleon) yields ^{11}B -rich boron ($^{11}\text{B}/^{10}\text{B} = 4.5$ or $\delta^{11}\text{B} = +110\text{‰}$)². Recent calculations by Cassé *et al.*¹⁰ show that the explosion of massive stars within giant molecular clouds could even yield $^{11}\text{B}/^{10}\text{B}$ ratios as high as 6.9. In Clayton's calculation, the variations in production rates between the different isotopes depend solely on the spallation cross-sections. Using this approach, the spallogenic B yield expected from these reactions was calculated (Table 1), with a collision rate Z between ^{12}C and H of 2×10^{39} atoms s^{-1} . Z is obtained from the observed γ -ray emission ($I_\gamma = 10^{-4}$ photons $\text{cm}^{-2} \text{s}^{-1}$), the distance ($d = 500$ pc) and the mass of Orion ($M = 10^5$ solar masses): $Z = 2/3 \times I_\gamma \times 4\pi d^2$ atoms s^{-1} (2/3 of the emission is attributed to ^{12}C).

In Table 1, the production rate of boron is proportional to cross-sections around 30 MeV and $\text{Si} = M/2.8 \times 10^4$ (2.8×10^4 is the H/Si cosmic ratio) with either 0 or 90% of Si retrieved from the gas by its condensation in solids. Calculated increases with time of B/Si and $\delta^{11}\text{B}$ values are compared in Table 2 and Fig. 2 with the present data, the Earth⁹ and the GCR³³. A progressive mixing with time between high-energy (GCR: $^{11}\text{B}/^{10}\text{B} = 2.5$, that is, $\delta^{11}\text{B} = -400\text{‰}$ and $\text{B/Si} = 10^{-5}$ or 10^{-6}) and low-energy spallogenic boron (referred to here as 'planetary boron' with $\delta^{11}\text{B} = +110\text{‰}$) is assumed. In such a model, because the isotopic compositions of the two endmembers depend solely on the energy of the spallogenic fluxes, the calculated differences in irradiation times does not imply a single star-formation region for ^{11}B -rich grains. From Fig. 1, it can be seen that meteoritic data are quite compatible with this model for durations of irradiation between 0.4 and 1 Myr which, in this respect, are consistent with those derived from extinct radioactive nuclides¹³. □

Received 14 October 1994; accepted 7 February 1995.

- Reeves, H. *Origin and Evolution of the Elements* (eds Prantzos, N., Vangioni-Flam, E. & Cassé, M.) 169–197 (Cambridge Univ. Press 1993).
- Meneguzzi, M., Audouze, J. & Reeves, H. *Astr. Astrophys.* **15**, 337–359 (1971).
- Mittler, H. E. *Astrophys. Space Sci.* **17**, 186–218 (1972).
- Reeves, H. A. *Rev. Astr. Astrophys.* **12**, 437–469 (1974).
- Prantzos, N., Cassé, M. & Vangioni-Flam, E. *Origin and Evolution of the Elements* (eds Prantzos, N., Vangioni-Flam, E. & Cassé, M.) 156–167 (Cambridge Univ. Press, 1993).

- Olive, K. A. & Schramm, D. N. *Nature* **360**, 439–442 (1992).
- Shima, M. *Geochim. cosmochim. Acta* **27**, 991–913 (1963).
- Agyei, E. K. & McMullen, C. C. U.S. Geol. Surv. Open-File Rep. **78-701**, 3–6 (1978).
- Chaussidon, M. & Jambon, A. *Earth planet. Sci. Lett.* **121**, 277–291 (1994).
- Cassé, M., Lehoucq, R. & Vangioni-Flam, E. *Nature* **373**, 318–319 (1995).
- Bloemen, H. *et al. Astr. Astrophys.* **281**, L5–L8 (1994).
- Anders, E. & Grevesse, N. *Geochim. cosmochim. Acta* **53**, 197–214 (1989).
- Lee, T., Papanastassiou, D. A. & Wasserburg, G. J. *Astrophys. J.* **211**, L107–L110 (1977).
- Podosek, F. A. *et al. Geochim. cosmochim. Acta* **55**, 1083–1110 (1991).
- McNaughton, N. J., Borthwick, S., Fallick, A. K. & Pillinger, C. T. *Nature* **294**, 639–641 (1981).
- Deloule, E. & Robert, F. *Geochim. cosmochim. Acta* (in the press).
- Ash, R. *Nature* **372**, 219–220 (1994).
- Alexander, C. M. O., Barber, D. J. & Hutchison, R. *Geochim. cosmochim. Acta* **53**, 3045–3057 (1989).
- Hewins, R. H. *Geochim. cosmochim. Acta* **55**, 935–942 (1991).
- Chaussidon, M. & Libourel, G. *Geochim. cosmochim. Acta* **57**, 5053–5062 (1993).
- Shaw, D. M., Higgins, M. D., Truscott, M. G. & Middleton, T. A. *Am. Mineral.* **73**, 894–900 (1988).
- Weller, M. R., Furst, M., Tombrello, T. A. & Burnett, D. S. *Geochim. cosmochim. Acta* **42**, 999–1009 (1978).
- Curtis, G., Gladney, E. & Jurney, E. *Geochim. cosmochim. Acta* **44**, 1945–1953 (1980).
- Shaw, D. M., Higgins, M. D., Hinton, R. W., Truscott, M. G. & Middleton, T. A. *Geochim. cosmochim. Acta* **52**, 2311–2319 (1988).
- Higgins, M. D. & Shaw, D. M. *Nature* **308**, 172–173 (1984).
- Curtis, D. B. & Gladney, E. S. *Earth planet. Sci. Lett.* **75**, 311–320 (1985).
- Davis, A. M., Hashimoto, A., Clayton, R. N. & Mayeda, T. K. *Nature* **347**, 655–658 (1990).
- Chaussidon, M. & Koeberl, C. *Geochim. cosmochim. Acta* **59**, 613–624 (1995).
- Kanzaki, T., Yoshida, M., Nomura, M., Kakihana, H. & Ozawa, T. *Geochim. cosmochim. Acta* **43**, 1859–1863 (1979).
- Nomura, M., Kanzaki, M., Ozawa, T., Okamoto, M. & Kakihana, H. *Geochim. cosmochim. Acta* **46**, 2403–2406 (1982).
- Bykov, A. & Bloemen, H. *Astr. Astrophys.* **283**, L1–L4 (1994).
- Clayton, D. D. *Nature* **368**, 222–224 (1994).
- Gibner, P. S., Mewaldt, R. A., Schindler, S. M., Stone, E. C. & Webber, W. R. *Astrophys. J.* **391**, L89–L92 (1992).

ACKNOWLEDGEMENTS. We thank A. G. W. Cameron and A. M. Davis for helpful reviews and G. J. Wasserburg and H. Reeves for their encouragement, scientific opinions and comments.

Chiral metal complexes with large octupolar optical nonlinearities

Christophe Dhenaut*, Isabelle Ledoux*,
Ifor D. W. Samuel*, Joseph Zyss*,
Manuel Bourgault† & Hubert Le Bozec†

* France Télécom CNET Centre Paris B, Laboratoire de Bagneux,
196, avenue Henri Rava, 92220 Bagneux, France
† Laboratoire de Chimie de Coordination Organique,
Université de Rennes, Avenue du Général Leclerc,
35042 Rennes, France

OPTICALLY nonlinear organic materials show considerable potential for applications in optical signal processing and telecommunications^{1,2}. Most materials are based on the *p*-nitro-aniline template, in which the optical nonlinearities are intimately associated with quasi-one-dimensional charge transfer. But there are problems associated with this conventional approach, arising from the strongly dipolar nature of the molecules². It has recently been recognized^{3–5} that two- and three-dimensional stereochemistry offers new possibilities for the design and synthesis of optically nonlinear molecules, in which charge transfer is multidirectional rather than dipolar in character; octupolar nonlinearities have now been demonstrated in several molecular systems^{5–7}. Tri-substituted ruthenium complexes⁶ appear particularly attractive because intense, multidirectional metal-to-ligand charge transfer leads to a significant enhancement of the optical nonlinearity, as quantified by the quadratic hyperpolarizability, β . Here we show that the choice of ligand can further increase β to values in excess of 10^{-27} e.s.u., comparable to the best dipolar optically nonlinear molecules.

The development of organic materials displaying high quadratic nonlinear optical susceptibilities (for frequency doubling, electro-optic modulation and related applications) depends on rational optimization of the underlying molecular structures.

Going beyond the classical dipolar approach for optimization of molecular hyperpolarizability, the concept of multipolar non-linearities was recently proposed, on the basis of group-theoretical and quantum-mechanical arguments³⁻⁵. Among the various octupolar molecules now available, ruthenium complexes, based on trigonal octupolar hybridization of a central metal ion by surrounding organic ligands⁶, are of special interest.

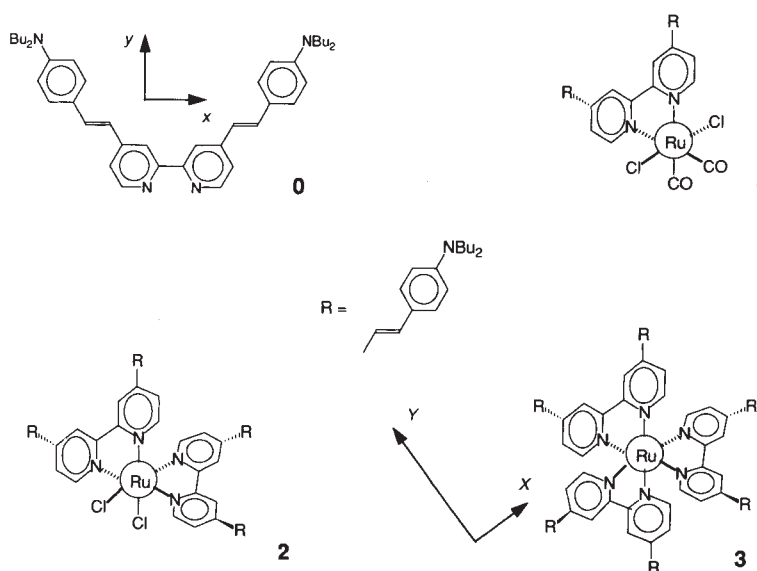
Preliminary investigations⁶ of Ru(II)-tris-bipyridine (RuTB) have found a significant octupolar static β value of the order of 50×10^{-30} e.s.u., in spite of the limited metal-to-ligand charge transfer (MLCT) between the metal and the unsubstituted ligand. Further substitution of the organic bipyridine ligands by *p*-dibutylaminostyryl groups is expected to enhance the β tensor, as compared to RuTB, owing to the extension of the conjugated π -electron system, and the introduction of a strongly electron-donating dimethylamino group, as compared to less active bipyridine moieties. The molecular structures are sketched in Fig. 1.

Ultraviolet-visible spectra of compounds 0-3 are displayed in Fig. 2. A large shift to the red is observed on metal complexation, especially for the tri-substituted derivative 3. The different behaviour of compound 2 is believed to relate to the presence of two isomers with different transition energies. These spectra clearly indicate a strong increase of the oscillator strength of the main charge-transfer transition when going from the pure ligand 0 to the tri-substituted complex 3.

In the case of purely octupolar (that is, dipole-free) and/or ionic molecules like compound 3, electric field poling and the related electric-field-induced second harmonic generation (EFISH)⁸ experiment are precluded, as they require uncharged species and a non-zero ground-state dipole. However, the hyperpolarizability β can be evaluated using the harmonic light scattering (HLS) technique^{9,10}. This method is based on the measurement of the second-harmonic contribution to Rayleigh scattering emitted by nonlinear molecules dissolved in a reference solvent. This method is based on an incoherent second-harmonic generation (SHG) process originating from fluctuations of optically induced microscopic dipoles, which can be observed in centrosymmetric media. The scattered SHG intensity $I_{2\omega}$ can be expressed as

$$I_{2\omega} = GN \sum_{ijklmn} \langle \beta_{ijk}(-2\omega; \omega, \omega) \beta_{lmn}^*(-2\omega; \omega, \omega) \rangle (I_\omega)^2 \quad (1)$$

FIG. 1 Molecular structures of the ligand and ruthenium complexes reported in this work. The (X, Y) frame is located in the mean plane of the propeller-shaped molecule 3, with Y along one of its two-fold axes. The (x, y) reference frame is attached to the planar organic ligand 0. The *trans* 4,4'-dibutylaminostyryl-2,2'-bipyridyl ligand 0 was obtained in 83% yield by dilithiation of 4,4'-dimethyl-2,2'-bipyridine with 2 equivalents of lithium diisopropylamide, followed by addition of *p*-dibutylaminobenzaldehyde and dehydration^{23,24}. The neutral complexes 1 and 2 were synthesized by reacting 0 with polymeric dicarbonyl dichlororuthenium ("ruthenium red carbonyl") in refluxing ethanol/water²⁵, and with ruthenium trichloride and lithium chloride in refluxing dimethylformamide²⁶, respectively. The homoleptic tris-bipyridine ruthenium complex 3 was prepared by reaction of the ligand 0 with ruthenium trichloride in refluxing dimethylformamide, and isolated as its PF₆⁻ salt by precipitation with an aqueous solution of ammonium hexafluorophosphate. All compounds exhibited spectroscopic and analytical data in accordance with assigned structures. Differential scanning calorimetry experiments did not give evidence of any decomposition of compound 3 below 330 °C. NMR studies performed on a 10^{-2} M solution of compound 3, irradiated at 350 nm by a set of 16 ultraviolet Hg lamps (power 16 W per lamp) for two hours show neither decomposition, nor any trace of free ligand or photo-isomerized species. According to ref. 17, the presence of the PF₆⁻ counterion strongly stabilizes the complex. Furthermore, after three hours of exposure of compound 3 to a powerful (10 times higher than in HLS experiments) 1.34 μ m Nd³⁺:YAG laser source, no change could be detected in the



Molecule	$\lambda_{\max}$ (nm)	$\langle \beta^2 \rangle^{1/2}$ (10^{-30} e.s.u.)	$\langle \beta_{\text{add}}^2 \rangle^{1/2}$ (10^{-30} e.s.u.)	$\langle \beta_{\text{add}}^2 \rangle^{1/2}$ (10^{-30} e.s.u.)
0	401	380	216	
1	467	420	183	
2	435, 606	>1,000	—	
3	513	2,200	740	312

Maximum-absorption wavelength $\lambda_{\max}$ (measured in dichloromethane), resonant $\langle \beta^2 \rangle^{1/2}$ (measured at 1.32 μ m) and non-resonant $\langle \beta_{\text{add}}^2 \rangle^{1/2}$ (deduced from $\langle \beta^2 \rangle^{1/2}$ using a degenerate three-level model, see ref. 12) values of compounds 0-3. (See Fig. 3 legend for a description of the $\langle \beta^2 \rangle^{1/2}$ nomenclature.) $\langle \beta_{\text{add}}^2 \rangle^{1/2}$ is the mean hyperpolarizability value calculated using an additive tensor model following equation (3). Relative experimental errors on β values are $\sim 15\%$. In the case of D₃ octupolar symmetry, and assuming isotropic averaging in the solution, $\langle \beta^2 \rangle^{1/2} = ((8/21)\beta_{\text{YYY}})^{1/2}$ and $\beta_{\text{YYY}} = -\beta_{\text{xxx}}$. This leads, in the case of compound 3, to a non-resonant β_{YYY} value of $1,200 \times 10^{-30}$ e.s.u. with $\beta_{\text{xxx}} = \beta_{\text{yxx}} = \beta_{\text{xyx}} = -1,200 \times 10^{-30}$ e.s.u. These 'off-diagonal' β tensor coefficients do not exist in polar one-dimensional systems. Results for compound 2 are only semiquantitative because this molecule is obtained in the form of two isomers, one of which is centrosymmetric.

where N is the molecular density, G is a factor incorporating local field corrections and geometrical parameters, I_ω is the intensity of the fundamental wave, and the β_{ijk} are the quadratic hyperpolarizability coefficients of individual molecules (0, 1, 2 or 3 in our case).

The averaging of the product of the cosine orientation factors connecting macroscopic (X, Y, Z) to microscopic (x, y, z) frames is performed, assuming isotropic distribution of the molecules and negligible intramolecular fluctuations. The result will depend on the symmetry of the molecule, and may be expressed as linear combinations of five basic quadratic polynomial expressions depending on β_{ijk} coefficients expressed in the molecular reference frame^{4,6,11}. Polarized detection at 2ω permits the inference of individual β_{ijk} tensor coefficients.

From the variation of the scattered harmonic emission with respect to a harmonic reference (that is, $(I_\omega)^2$), it is possible, from HLS measurements at various concentrations of the

ultraviolet-visible spectrum of the molecule; this confirms that no detectable multiphoton isomerization or decomposition has taken place in our experiments.

unknown molecule—as sketched in Fig. 3 for compound 3—to infer $\langle \beta_{ijk}(-2\omega; \omega, \omega) \beta_{lmn}^*(-2\omega; \omega, \omega) \rangle$. Experimental results are reported in Table 1. Remarkably high β values (up to 2×10^{-27} e.s.u.) are observed for compounds 2 and 3, corresponding to a ten-fold enhancement relative to the unsubstituted RuTB molecule. Resonance enhancement is partly responsible for these large β s. According to the three-level quantum model with a ground state $|0\rangle$ and a degenerate set of eigenstates $|1\rangle$ and $|1'\rangle$ to account for β -frequency dispersion¹², static $\beta(0)$ values can be easily estimated from experimental data as follows

$$\beta = \frac{3}{2} \frac{|\mu_{01}^2| |\mu_{1\bar{1}}|}{E_{01}^2} \frac{E_{01}^4}{(E_{01}^2 - 4\hbar\omega^2)(E_{01}^2 - \hbar\omega^2)} \\ = \beta(0) \frac{E_{01}^4}{(E_{01}^2 - 4\hbar\omega^2)(E_{01}^2 - \hbar\omega^2)} \quad (2)$$

where $\mu_{1\bar{1}}$ (or μ_{01}) is the transition dipole moment connecting the degenerate excited state $|1\rangle$ to its counterpart $|\bar{1}\rangle$ (or to the ground state $|0\rangle$), and E_{01} is the energy difference between the degenerate excited state and the ground state. An adequate comparison of the nonlinear responses of the molecules can be undertaken on the basis of the static $\beta(0)$ values only, which reflect their intrinsic hyperpolarizabilities, regardless of the energy detuning between harmonic (or fundamental) photon energies and that of the degenerate excited states. For compound 3, $\beta(0)$ approaches 10^{-27} e.s.u., comparable to the record value of long dipolar donor-acceptor conjugated molecules¹³⁻¹⁶, and with the crucial additional advantage of the excellent thermal and photochemical stability of ruthenium complexes¹⁷.

Whereas no significant β enhancement is noticed in monosubstituted complex 1 as compared to the pure ligand 0—contrary to previous measurements of the vector part β_{vec} of the hyperpolarizability tensor using the EFISH technique applied to strongly one-dimensional complexes^{18,19}—a large enhancement is observed for compounds 2 and 3. It is instructive to compare these β values to those obtained from a purely additive model, where the nonlinear response of the complex is assumed to result from the tensorial sum of the individual responses of the three

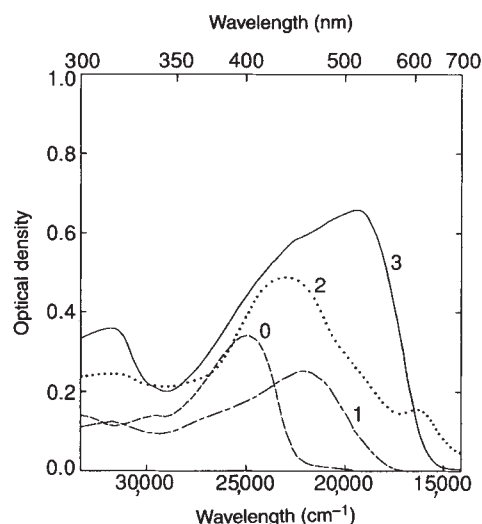


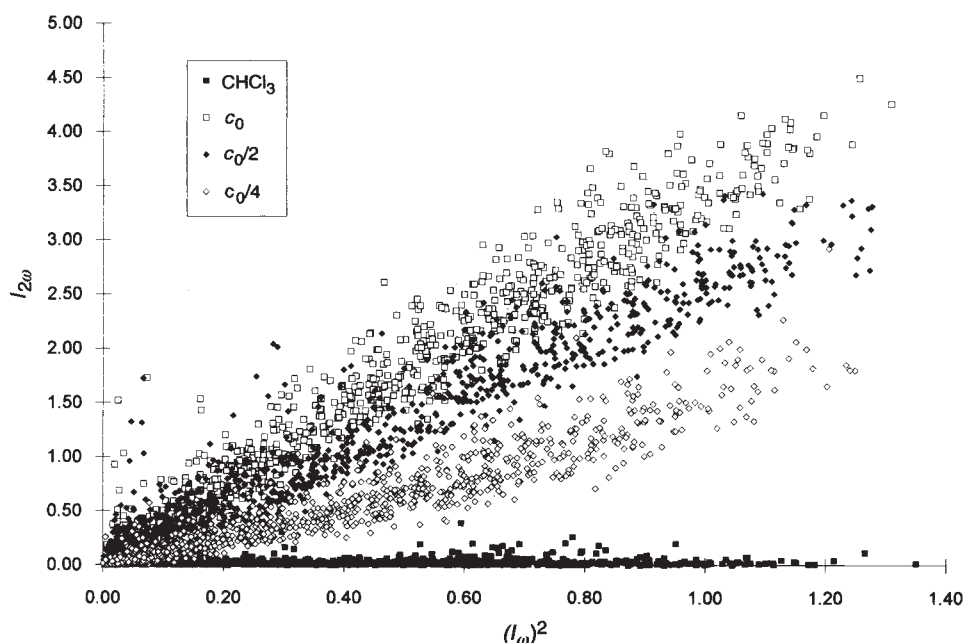
FIG. 2 Ultraviolet-visible absorption spectra of compounds 0-3 (Fig. 1). The concentrations are 1.2×10^{-5} M for compound 0, and 10^{-5} M for compounds 1, 2 and 3.

ligand subunits arranged in the same propeller geometry as in compound 3, but without the ruthenium atom. The additive value is given by

$$\beta_{YY}^{\text{add}} = \frac{3}{4} (\beta_{yyy} - 3\beta_{yxx} \cos^2 \theta) \quad (3)$$

where β_{yxx} and β_{yyy} are the hyperpolarizabilities of the pure ligand. The coordinates x, y, z (or X, Y, Z) refer to the cartesian axes of the ligand (or of the complex 3), y being the two-fold axis of the ligand, as shown in Fig. 1. θ is the angle between the xy molecular plane of the ligand and the XY molecular plane of compound 3, assumed to be 60° . The individual β_{yxx} and β_{yyy} coefficients are calculated from measurement of the HLS depolarization factor of compound 0. It is preferable to compare the static $\beta(0)$ values of the pure ligand and metal complexes

FIG. 3 Experimental dependence of the second-harmonic intensity $I_{2\omega}$ as a function of the square of the fundamental intensity $(I_\omega)^2$ for solutions of compound 3 at various concentrations ($c_0 = 1.06 \times 10^{-3}$ M) and chloroform, which is taken as a reference. The experiments are performed at a fundamental wavelength of $1.32 \mu\text{m}$, generated by a Q-switched, mode-locked Nd:YAG laser. This wavelength, being red-shifted with respect to the more widely used $1.06 \mu\text{m}$ emission from Nd:YAG lasers, ensures that the second harmonic is away from the absorption bands of the compounds under investigation, as shown by the ultraviolet-visible absorption spectra in Fig. 2 (except for compound 2, for which very low concentrations, of the order of 10^{-15} molecules per cm^3 , are used to minimize the absorption of the second harmonic at 660 nm). Moreover, it also precludes significant hot two-photon fluorescence contributions, and reduces the effects of resonance. The experiment consists in recording the variation of the second-harmonic signal with respect to a "reference" fundamental (which is proportional to $(I_\omega)^2$) for several solutions of increasing concentrations. Plotting the slopes of the resulting lines as a function of these concentrations enables direct



determination of $\langle \beta_{ijk}(-2\omega; \omega, \omega) \beta_{lmn}^*(-2\omega; \omega, \omega) \rangle$. In the text, " β " values correspond to $\langle \beta^2 \rangle^{1/2}$.

in view of the different resonance enhancements of the various systems. Static β_{yyy} and β_{yxx} values are 452×10^{-30} and -300×10^{-30} e.s.u., respectively.

The experimental value of β for compound 3 is more than double that deduced by the additive model (that is, $\beta_{yyy}^{\text{add}} = 500 \times 10^{-30}$ e.s.u. compared to $1,200 \times 10^{-30}$ e.s.u. for the corresponding experimental value). This clearly indicates that the ruthenium atom in an octupolar geometry introduces an additional charge-transfer contribution to β on complexation of the three ligands by the metal, leading to an important enhancement of the nonlinearity as compared to that of the corresponding purely organic systems, as has been observed for more conventional dipolar species^{18,19}. These results illustrate also the interest of three-dimensional organometallic structures as compared to rod-like geometries: enhancement of the octupolar part of the β tensor has led to β values in a magnitude range which has been previously believed to be unreachable for compounds other

than one-dimensional dipolar systems. Moreover, large off-diagonal β_{yxx} , β_{xyx} and β_{xxy} tensor coefficients, which are equal to $-\beta_{yyy}$ in the case of octupolar molecules, are not available in the case of a one-dimensional tensor. Such multidirectional charge-transfer processes cooperate to enhance strongly the β tensor.

A current challenge is the macroscopic organization of non-linear octupoles in non-centrosymmetric assemblies, remembering that electric field poling is precluded by the absence of permanent dipole moments. However, combined one-photon absorption at 2ω and two-photon absorption at ω (where ω is the fundamental frequency of a laser) will break the centrosymmetric order of the irradiated medium²⁰, resulting in a periodic absorption grating with quasi-phase-matching potential. The resulting 'octupolar' ordering can be 'frozen' by associating the combined absorption processes with reversible photoinduced effects^{21,22}. □

Received 13 June 1994; accepted 15 February 1995.

1. Chemla, D. S. & Zyss, J. (eds) *Nonlinear Optical Properties of Organic Molecules and Crystals* (Academic, Boston, 1987).
2. Zyss, J. (ed.) *Molecular Nonlinear Optics: Materials, Physics and Devices* (Academic, Boston, 1993).
3. Zyss, J. *Nonlin. Opt.* **1**, 3–18 (1991).
4. Zyss, J. *J. chem. Phys.* **98**, 6583–6599 (1993).
5. Zyss, J. & Ledoux, I. *Chem. Rev.* **94**, 77–105 (1994).
6. Zyss, J., Dhenaut, C., Chau Van, T. & Ledoux, I. *Chem. Phys. Lett.* **206**, 409–414 (1993).
7. Verbiest, T. et al. *Opt. Lett.* **18**, 525 (1993).
8. Oudar, J.-L. *J. chem. Phys.* **67**, 446–457 (1977).
9. Terhune, R. W., Maker, P. D. & Savage, C. M. *Phys. Rev. Lett.* **14**, 681–684 (1965).
10. Maker, P. D. *Phys. Rev. A* **1**, 923–951 (1970).
11. Cyvin, S. G., Rauch, J. E. & Decluis, J. C. *J. chem. Phys.* **43**, 4083–4095 (1965).
12. Joffe, M., Yaron, D., Silbey, R. J. & Zyss, J. *J. chem. Phys.* **97**, 5607–5615 (1992).
13. Barzoukas, M., Blanchard-Desce, M., Josse, D., Lehn, J.-M. & Zyss, J. *Chem. Phys.* **133**, 323–329 (1989).

14. Blanchard-Desce, M., Lehn, J.-M., Barzoukas, M., Ledoux, I. & Zyss, J. *Chem. Phys.* **181**, 281–289 (1994).
15. Rao, V. P., Jen, A. K. Y., Wong, K. Y. & Drost, K. J. *J. chem. Soc., chem. Commun.* 1118–1120 (1993).
16. Gilmour, S. et al. *Chem. Mater.* **6**, 1603–1604 (1994).
17. Juris, A. et al. *Coord. Chem. Rev.* **84**, 85–277 (1988).
18. Loucif-Saïbi, R. et al. *Chem. Phys.* **167**, 369–375 (1992).
19. Bourgault, M. et al. *J. chem. Soc. chem. Commun.* 1623–1624 (1993).
20. Nunzi, J.-M., Charra, F., Fiorini, C. & Zyss, J. *Chem. Phys. Lett.* **219**, 349–354 (1994).
21. Charra, F., Kajzar, F., Nunzi, J.-M., Raimond, P. & Idiat, E. *Opt. Lett.* **18**, 941–943 (1993).
22. Fiorini, C., Charra, F., Nunzi, J.-M. & Zyss, J. in *Proc. Eur. Conf. Laser and Electro-optics* 33–34 (IEEE, Piscataway, 1994).
23. Juris, A., Campagna, S., Bidd, I., Lehn, J.-M. & Ziessel, R. *Inorg. Chem.* **27**, 4007–4011 (1988).
24. Kocian, O., Mortimer, R. J. & Beer, P. D. *J. chem. Soc., Perkin Trans. I* 3203–3205 (1990).
25. Kelly, J. M., O'Connell, C. M. & Vos, J. G. *Inorg. Chim. Acta* **64**, L75–L76 (1982).
26. Sullivan, B. P., Salmon, D. J. & Meyer, T. J. *Inorg. Chem.* **17**, 3334–3341 (1978).

Prediction of crystal growth morphology based on structural analysis of the solid–fluid interface

X. Y. Liu*, E. S. Boek†, W. J. Briels‡ & P. Bennema*

* Department of Solid State Chemistry, University of Nijmegen, Toernooiveld 1, 6525 ED Nijmegen, The Netherlands

† Schlumberger Cambridge Research, High Cross, Madingley Road, Cambridge CB3 0EL, UK

‡ Chemical Physics Laboratory, University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands

PREDICTING the shape of growing crystals is important for industrial crystallization processes. The equilibrium form of a crystal can be determined unambiguously from a consideration of the surface free energies of the various crystallographic faces $\{hkl\}$ ¹, but the growth morphology is determined by kinetic factors which are harder to predict. This morphology depends on the relative growth rates R_{hkl}^{rel} of the crystal faces. Several theories have been advanced^{2,3} to relate R_{hkl}^{rel} to geometric or energetic characteristics of the surfaces $\{hkl\}$, but these have met with limited success in predicting the crystal morphologies observed. Here we present a theoretical approach to the problem in which R_{hkl}^{rel} is determined by quantities that are accessible either from kinetic models or from computer simulations of the solid–fluid interface. The important parameters controlling the growth rate are the energy required to create a step at the crystal surface and the free-energy barrier for an adsorbed solute molecule to be incorporated into the crystal. Both can be related to the mole fraction of adsorbed solute molecules in dynamic equilibrium with those in the crystal surface.

When this approach is applied to the case of urea crystals grown from aqueous solution, we predict a needle-like shape which is consistent with experimental observations.

The growth of crystals bounded by facets usually occurs in a spiral fashion^{4–7}, whereby one or more screw dislocations on the crystal surface give rise to spiral steps (Fig. 1). Growth takes place by the deposition and incorporation of solute molecules from the bulk onto these steps, causing them to advance with a speed v_{step} along the surface, perpendicularly to the step (Fig. 1a, b). Denoting the average distance between two steps as λ_0 and their heights as d_{hkl} , the growth rate of the surface is given by $R_{hkl} = v_{\text{step}} d_{hkl} / \lambda_0$ (Fig. 1b). In the case of crystals grown from solutions, the delivery of molecules at the steps is governed by the diffusion of solute molecules from solution to the kinks⁴ (Fig. 1c). In this process, the incorporation of these 'growth units' into kinks is rate-determining at relatively low supersaturations $\Delta\mu/kT$, and $v_{\text{step}} \propto X_{A(hkl)} \rho_{hkl}^{\text{kink}} \exp \{-\Delta G_{hkl}^{\ddagger}/kT\}$ for a given $\Delta\mu/kT$ (Fig. 1c). It follows that the relative growth rate of crystal faces $\{hkl\}$ is

$$R_{hkl}^{\text{rel}} \propto X_{A(hkl)} d_{hkl} \rho_{hkl}^{\text{kink}} \exp \{-\Delta G_{hkl}^{\ddagger}/kT\} / \lambda_0 \quad (1)$$

where ρ_{hkl}^{kink} is the kink density, $X_{A(hkl)}$ is the solute concentration at the face $\{hkl\}$, and $\Delta G_{hkl}^{\ddagger}$ is the activation free energy for adsorbed solute molecules to be incorporated into kinks. To relate the relative growth rate R_{hkl}^{rel} to quantities that can be easily obtained from a molecular-dynamics simulation, we introduce the concept of 'effective growth units'⁸, which are those solute molecules adsorbed at the surface that are in dynamic equilibrium with the solid molecules at the kinks (called S^{I} molecules). All adsorbed solute molecules may be classified into two classes, F_1 and F_2 . The former are those molecules that have roughly the same orientation or conformation as solid molecules at the surface (Fig. 3). To be incorporated into a kink, an F_1 -molecule needs to surpass a barrier $\Delta G_{\text{kink}}^{\ddagger}$, which corresponds to desolvation free energy. F_2 molecules are all other adsorbed solute molecules. Before they can be incorporated into

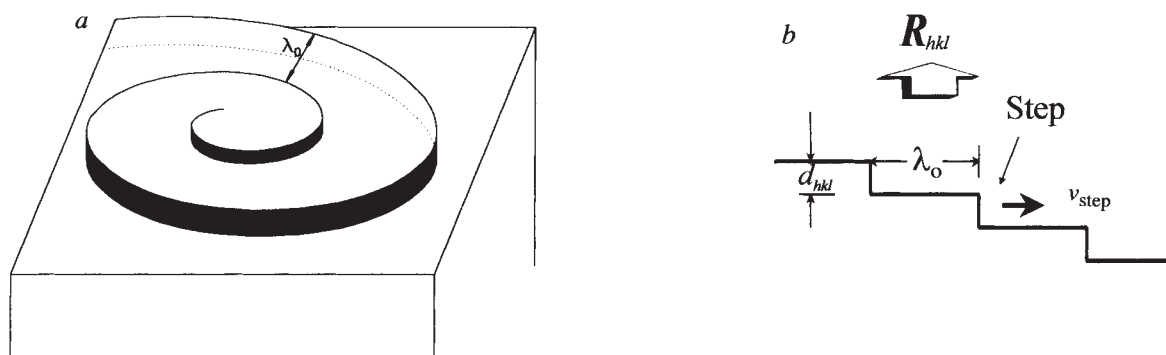


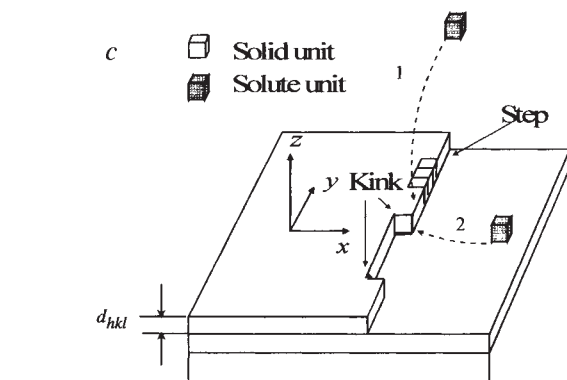
FIG. 1 a, Growth spiral for crystal surface growing from solution. The spiral is rotating with constant angular velocity under constant supersaturation $\Delta\mu/kT$ ($\Delta\mu = \mu^f - \mu^s$, where μ^f is the chemical potential of a solute molecule and μ^s of the solid molecule). The spiral is guided by the screw dislocation centre. Some distance from the centre, the spiral becomes roughly equal to concentric circular steps with a separation $\lambda_0 \approx 19r_c$ (refs 4–7). Here r_c is the radius of the two-dimensional critical nucleus given by $r_c \approx \phi_{\text{step}} \Omega / \Delta\mu$, ϕ_{step} is the average step energy of a step per growth unit (compare with c) and Ω is the molecular volume. Therefore, $\lambda_0 \propto \phi_{\text{step}}$ at a given $\Delta\mu/kT$. b, From a two-dimensional cut through a, we see that the growth rate of a face (hkl), R_{hkl} is equal to the flux of steps given by $F_{\text{step}} = v_{\text{step}} \rho_{\text{step}} = v_{\text{step}} / \lambda_0$ times d_{hkl} (ρ_{step} is step density). So $R_{hkl} = d_{hkl} v_{\text{step}} / \lambda_0$. c, Surface model showing steps and kinks. A growth unit can be incorporated into the crystal structure at kink sites. The shape of v_{step} depends on the means of delivery of growth units to the steps⁴. For the growth of crystals from a solution, the surface diffusion of solute molecules to steps (route 2) is irrelevant, because the mean square displacement of an adsorbed solute molecule at the surface is short. Solute molecules are delivered directly from the bulk to steps at the surface by volume diffusion (route 1). This includes two main steps: the delivery of growth units to the steps by means of volume diffusion, and the integration of growth molecules into the crystal at kinks. At relatively low supersaturations, the latter is the rate determining step, and the rate is proportional to the concentration of adsorbed solute molecules $X_{A(hkl)}$ around the kinks, ρ_{hkl}^{kink} and $\exp\{-\Delta G_{hkl}^{\ddagger}/kT\}$. Therefore, $v_{\text{step}} \propto X_{A(hkl)} \rho_{hkl}^{\text{kink}} \exp\{-\Delta G_{hkl}^{\ddagger}/kT\}$ for a

a kink, F_2 molecules must be transformed into F_1 molecules, for which they have to surpass a barrier ΔG^* . Effective growth units consist of F_1 molecules and those F_2 molecules for which ΔG^* is negligibly small. Two typical cases can be distinguished: $\Delta G_{\text{kink}}^* > \Delta G^*$ (Fig. 2a). In this case both F_1 and F_2 molecules have equal probability of being incorporated into the kinks, and both are effective growth units. The mole fraction of effective growth units is $X_{A(hkl)}^{\text{eff}} \approx X_{A(hkl)} = X_{F_1} + X_{F_2}$. $\Delta G_{\text{kink}}^* < \Delta G^*$ (Fig. 2b). In this case, F_2 molecules cannot be easily incorporated into the kinks and are not effective growth units. From the point of view of equilibrium thermodynamics, F_2 molecules may be said to be a separate species; in other words, the mole fraction of effective growth units is $X_{A(hkl)}^{\text{eff}} \approx X_{F_1}$.

Now we return to equation (1). To be incorporated into the kinks, adsorbed solute molecules should have the right orientation with respect to the solid molecules at the surface (refs 4, 5; X.Y.L. and P.B., manuscript in preparation). Given the definition of effective growth units, the probability for adsorbed solute molecules to obtain the proper orientation is about equal to $X_{A(hkl)}^{\text{eff}} / X_{A(hkl)}$, leading to the relation $\exp\{-\Delta G_{hkl}^{\ddagger}/kT\} \approx (X_{A(hkl)}^{\text{eff}} / X_{A(hkl)}) \exp\{-\Delta G_{hkl}^*/kT\}$. ΔG_{hkl}^* is roughly independent of the specific surface^{3,4}, and can be considered a constant in this relationship.

The average separation of steps λ_0 is roughly proportional to the average step energy per molecule ϕ_{hkl}^{step} (see Fig. 1a legend), and $\rho_{hkl}^{\text{kink}} \approx \exp\{-\phi_{hkl}^{\text{kink}}/kT\}$ (refs 4–7). Because the average kink energy $\phi_{hkl}^{\text{kink}} \approx \phi_{hkl}^{\text{step}}$ (see Fig. 1c legend), $\rho_{hkl}^{\text{kink}} \approx \exp\{-\phi_{hkl}^{\text{step}}/kT\}$. Combining the results so far, we find

$$R_{hkl}^{\text{rel}} \propto d_{hkl} X_{A(hkl)}^{\text{eff}} \exp\{-\phi_{hkl}^{\text{step}}/kT\} / \phi_{hkl}^{\text{step}} \quad (2)$$



given ($\Delta\mu/kT$) (ref. 4). The step energy ϕ_{step} (per structural unit) corresponds to the energy cost to create a step of a structural unit length. The kink energy ϕ_{kink} is then the energy cost to create a kink at a step. Assume that the step energy in the x-direction (shown here) is ϕ_x^{step} , and in the y-direction ϕ_y^{step} . It can easily be seen that ϕ_{kink} in the x-direction step is roughly equal to ϕ_x^{step} , and in the y-direction step is ϕ_y^{step} . If $\phi_x^{\text{step}} = \phi_y^{\text{step}}$, $\phi_{\text{step}} \approx \phi_{\text{kink}}$. Otherwise, the average step energy ϕ_{step} is equal to the average kink energy ϕ_{kink} .

Obviously, to find the expression of ϕ_{hkl}^{step} is one of the key issues. In the following, we shall arrive at

$$\phi_{hkl}^{\text{step}} \approx \xi_{hkl} (\ln X_{A(hkl)}^{\text{eff}} / \ln X_A) \Delta H_{hkl}^{\text{diss}} / n_{hkl} \quad (3)$$

Here n_{hkl} is the coordination number, that is, the number of neighbours of a solid molecule within the two-dimensional crystal slice (hkl), X_A is the concentration of solute molecules in the bulk, and $\Delta H_{hkl}^{\text{diss}}$ is the experimentally observed enthalpy of dissolution. The crystallographic orientation factor ξ_{hkl} will be defined below.

The step energy, the one-dimensional energy cost of creating a step with the length of one molecule, can be related to the three-dimensional local dissolution enthalpy $\Delta H_{hkl}^{\text{diss}}$ at the crystal surface (ref. 9; X.Y.L. and P.B., manuscript in preparation), which is the enthalpy change due to transforming a solid molecule at a kink to an adsorbed solute molecule. We may reduce $\Delta H_{hkl}^{\text{diss}}$ to a sum of in-plane contributions by multiplying it by $\xi_{hkl} = E_{hkl}^{\text{slice}} / E^{\text{cr}}$, where E^{cr} is the lattice energy per molecule and E_{hkl}^{slice} is the two-dimensional lattice energy per molecule for a crystal slice (hkl) with thickness d_{hkl} (ref. 3). We may approximate $\phi_{hkl}^{\text{step}} n_{hkl} \approx \xi_{hkl} \times \Delta H_{hkl}^{\text{diss}}$. We notice that because of the ordering of fluid molecules at the interface and the crystal relaxation near the surface, $\Delta H_{hkl}^{\text{diss}}$ is normally different from ΔH^{diss} (ref. 9). Nevertheless, we may still easily estimate $\Delta H_{hkl}^{\text{diss}} / \Delta H^{\text{diss}}$ (ref. 9). From the van't Hoff relation¹⁰ we have $\ln X_A = -\Delta H^{\text{diss}}/kT + \Delta H^{\text{m}}/kT^{\text{m}} \approx \Delta H^{\text{diss}}(T - T^{\text{m}})/kTT^{\text{m}}$ (the superscript m denotes melting). Applying the same theories to the dissolution process at the crystal surface⁹, we have $\ln X_{A(hkl)}^{\text{eff}} = -\Delta H_{hkl}^{\text{diss}}/kT + \Delta H_{hkl}^{\text{m}}/kT^{\text{m}} \approx \Delta H_{hkl}^{\text{diss}}(T - T^{\text{m}})/kTT^{\text{m}}$. From the above relations we get $\Delta H_{hkl}^{\text{diss}} / \Delta H^{\text{diss}} \approx$

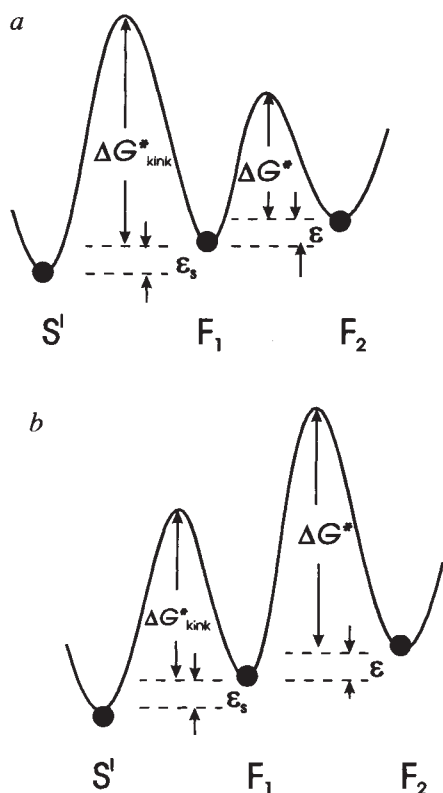
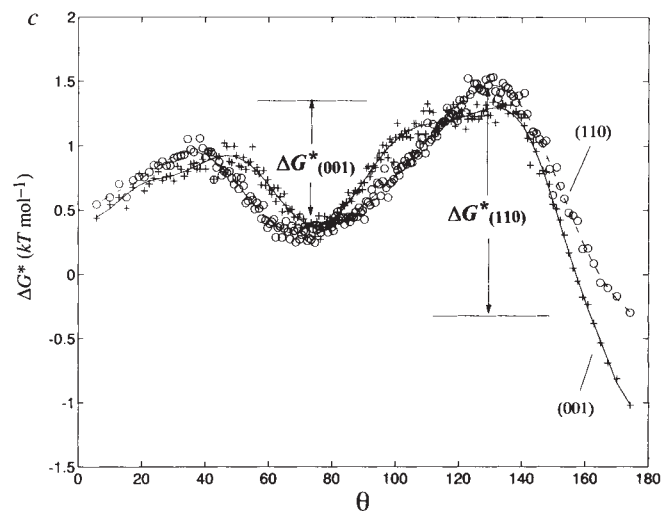


FIG. 2 *a* and *b*, The potential landscape between different types of structural units at the surface. In equilibrium with interfacial solid molecules S' , F_1 molecules change directly into S' molecules by passing the potential barrier ΔG^*_{kink} , while in a similar process, F_2 molecules should first change to F_1 , and then to S' molecules. (ΔG^* is the potential barrier for the transition from F_2 to F_1 ; $\epsilon_s = \mu_{F_1}^0 - \mu_{S'}^0$; $\epsilon = \mu_{F_2}^0 - \mu_{F_1}^0$; μ^0 is the standard chemical potential for a certain species.) Thus, the rates of transition from a pure F_1 to an S' molecule and from a pure F_2 to a pure F_1 molecule are given by $R^{I-S} = \nu_{F_1} \exp(-\Delta G^*_{\text{kink}}/kT)$ and $R^{II-I} = \nu_{F_2} \exp(-\Delta G^*/kT)$, respectively. Here ν_q denotes the frequency of thermal vibration of species q . So the rate of the transition from a pure F_2 to an S' molecule is $R^{II-S} = R^{I-S} R^{II-I} / (R^{II-I} + R^{I-S})$. It is shown in *a* that $\Delta G^*_{\text{kink}} > \Delta G^*$, implying that $R^{II-I} \gg R^{I-S}$ and $R^{II-S} \approx R^{I-S}$. In this case both F_1 and F_2 molecules can be treated as effective growth units. In *b*, it is shown that $\Delta G^*_{\text{kink}} < \Delta G^*$, implying that $R^{II-I} \ll R^{I-S}$ and $R^{II-S} \ll R^{I-S}$. In this case only F_1 molecules can be considered as the effective growth units.



c, The potential of adsorbed urea molecules plotted against orientations of the dipole vectors of the molecules at the (001) interface (indicated by +) and the (110) interface (indicated by O). The potential can be expressed as a function of the orientational distribution probability $P(\cos \theta)$ by $-kT \ln [P(\cos \theta)]$ (θ is the angle between the dipole vectors and the outward surface normal). The distribution probability $P(\cos \theta)$ has been calculated by averaging the time series of coordinates of the solute urea molecules in the first adsorbed liquid layer. For the orientation of (001) a small minimum at $\theta \approx 0$ in our present calculations correspond to the ordering of urea molecules occurring in the second fluid layer. The face (001) corresponds to case I (shown in *a*) and the face (110) corresponds to case II (shown in *b*). To determine the orientations of urea molecules fully, we should also look at the distribution of the N-N vector of urea molecules; similar analysis led to an extra factor of 0.5 in $X_{\text{A}(hkl)}^{\text{eff}}$ for {110} and 1 for {001}.

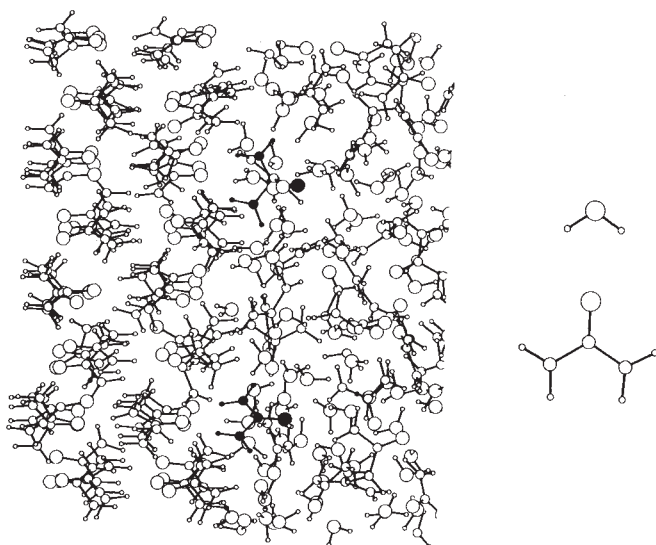


FIG. 3 Snapshot of the interface between crystalline urea (left) and saturated aqueous urea solutions (right) in the {001} orientations. On the right, a water molecule (above) and a urea molecule (below) are drawn for clarification. These snapshots have been generated by means of molecular dynamics simulations¹⁵⁻¹⁷. Using this technique, the newtonian equations of motion are integrated numerically for a large number of particles. The computational box consisted of 300 crystalline urea molecules, 200 solute urea molecules and 500 water molecules. Only the interfacial region of the computational box is shown. Simulation runs of about 200 picoseconds have been done for both interfaces. By averaging the time series of the generated coordinates, density profiles and orientational distributions of the adsorbed solute urea molecules at the interface have been calculated. Two F_1 -like urea molecules in the first fluid layer of the (001) interface are drawn in black. (The intermolecular potentials chosen for the calculation are given in refs 15-17, and the self-consistency of the calculations has been checked by comparing the calculated total nitrogen radial distribution function with that of neutron scattering experiments¹⁷. The results are in good agreement.)

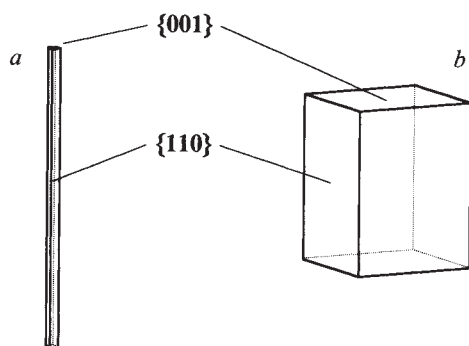


FIG. 4 Wulff constructions of the growth morphology of urea crystals based on different formalisms. *a*, The growth morphology predicted according to our approach (see equation (2)) ($R_{001}^{\text{rel}}/R_{110}^{\text{rel}} \approx 34.9$). *b*, The growth morphology predicted according to the Hartman–Perdok theory^{3,11,18}, where it was assumed that $R_{hkl}^{\text{rel}} \approx (1 - \xi_{hkl})$ ($R_{001}^{\text{rel}}/R_{110}^{\text{rel}} \approx 1.35$). The growth habits of urea crystals shown in *a* and *b* are substantially different. One is needle-like, the other is block-like. Experiments^{13,14} show that urea crystals grown from aqueous solutions have a needle-like shape ($R_{001}^{\text{rel}}/R_{110}^{\text{rel}} \approx 50$). $d_{001} = 4.712 \text{ \AA}$, $d_{110} = 4.003 \text{ \AA}$ (ref. 12); $\xi_{001} = 0.556$, $\xi_{110} = 0.671$ (ref. 18); $n_{001} \approx n_{110} \approx 4$ (ref. 18); $X_{A(001)}^{\text{eff}} \approx 0.031$, $X_{A(110)}^{\text{eff}} \approx 0.088$ (this work); $X_A = 0.256$ (ref. 17), and $\Delta H^{\text{diss}}/kT = 6.79$ ($T = 298.15 \text{ K}$) (ref. 19).

In $X_{A(hkl)}^{\text{eff}}/\ln X_A$ (ref. 9), from which we finally obtain equation (3). Thus, the growth rate of a specific crystal face can be determined from equations (2) and (3). The quantities n_{hkl} , d_{hkl} and ξ_{hkl} can easily be calculated from the crystal structure and a given potential model (refs 3, 11 and X.Y.L. and P.B., manuscript in preparation). The key issue then is to calculate $X_{A(hkl)}^{\text{eff}}$ from an analysis of the interfacial structure obtained from molecular dynamics simulations.

We now apply the above theory to urea crystals grown from aqueous solutions. Urea ($\text{O}=\text{C}(\text{NH}_2)_2$) crystallizes in the tetragonal space group $P_{42/m}$ ($a = 5.661$, $c = 4.712 \text{ \AA}$)¹². According to experiments^{13,14} the crystals are mainly bounded by the {001}, the {110} and the {111} faces. We shall ignore the {111} faces for the sake of simplification. To calculate $X_{A(hkl)}^{\text{eff}}$ and $X_{A(hkl)}$ for the remaining two surfaces, we have carried out molecular dynamics simulations^{15–17} for the two interfaces between a saturated urea solution and the crystal surfaces, at $X_A = 0.256$. A snapshot along the interface of {001} is shown in Fig. 3. Potential distributions of the dipole vectors ($\text{O} \rightarrow \text{C}$) of solute urea at the surfaces of {001} and {110} are shown in Fig. 2c.

F_1 urea molecules in the first fluid layer at the {001} surfaces should have their $\text{O} \rightarrow \text{C}$ dipoles roughly antiparallel to the outward surface normal ($\theta \approx 180^\circ$), where θ is the angle between the vectors and the outward surface normal; see Fig. 3). In Fig. 2c, an eminent minimum occurs at $\theta \approx 180^\circ$. The transformation from other orientations to the orientation of $\theta \approx 180^\circ$ may need to pass at most the small potential barrier $\Delta G_{(001)}^*$ ($\leq 1 kT$). It can be roughly estimated⁴ that $\Delta G_{\text{kink}}^* \approx kT$. This suggests that ΔG^* is small compared with ΔG_{kink}^* , and we have the case where $X_{A(001)}^{\text{eff}} \approx X_{A(001)}$.

Unlike the {001} faces, F_1 -like urea molecules at the {110} faces should have their $\text{O} \rightarrow \text{C}$ dipoles roughly parallel to the surface ($\theta \approx 90^\circ$)^{15,18}. It is shown in Fig. 2c that $\Delta G_{(110)}^*$ is relatively high ($\Delta G_{(110)}^* \approx 2\Delta G_{(001)}^*$), so here we have the second case. Only those urea molecules with $60^\circ \leq \theta \leq 120^\circ$ are regarded as F_1 -like molecules or effective growth units. (ΔG^* of those molecules are equal to or less than zero; see Fig. 2c). From $X_{A(110)}$ and Fig. 2c, $X_{A(110)}^{\text{eff}}$ can be calculated.

Substituting these data, together with n_{hkl} , d_{hkl} and ξ_{hkl} (refs 12, 18), into equations (2) and (3) yields R_{hkl}^{rel} for the two orientations (see Fig. 4 legend for details). The growth form of urea crystals is constructed based on R_{hkl}^{rel} in corresponding crystallographic orientations^{1,3}. We predict that, in the case of

growth from aqueous solution, urea crystals should adopt a needle-like shape (Fig. 4a), in agreement with experimental observations^{12,13}. This striking result demonstrates the validity of our approach in the prediction and the study of morphology of crystals growing from solutions. □

Received 23 September 1994; accepted 7 February 1995.

- Kern, R. in *Morphology of Crystals* (ed. Sunagawa, I.) 77–206 (Terra, Tokyo, 1987).
- Donnay, J. D. H. & Harker, D. *Am. Mineral.* **22**, 446–447 (1937).
- Hartman, P. in *Morphology of Crystals* (ed. Sunagawa, I.) 269–319 (Terra, Tokyo, 1987).
- Chernov, A. A. *Modern Crystallography III—Crystal Growth* 1–158 (Springer, Berlin, 1984).
- Bennema, P. & Gilmer, J. in *Crystals Growth: An Introduction* (ed. Hartman, P.) 263–327 (North Holland, Amsterdam, 1973).
- Burton, W. K., Cabrera, N. & Frank, F. C. *Phil. Trans. R. Soc.* **243**, 299–358 (1951).
- Ohara, M. & Reid, R. C. *Modeling Crystal Growth Rates from Solution* 1–165 (Prentice-Hall, Englewood Cliffs, 1973).
- Liu, X. Y. & Bennema, P. *Phys. Rev.* **B49**, 765 (1994).
- Liu, X. Y. & Bennema, P. *J. chem. Phys.* **98**, 5863–5875 (1993); Liu, X. Y. & Bennema, P. in *Current Topics in Crystal Growth Research* (in the press).
- Fowler, S. R. & Guggenheim, E. A. *Statistical Thermodynamics* (Cambridge, London, 1960).
- Bennema, P. in *Handbook on Cryst. Growth* (ed. Hurler, D. T. J.) 477–581 (North-Holland, Amsterdam, 1993).
- Worsham, J. E., Levy, J. H. A. & Peterson, S. W. *Acta cryst.* **10**, 319 (1957).
- Davey, R., Fila, W. & Garside, J. *J. Crystal Growth* **79**, 607–613 (1986).
- Docherty, R., Roberts, K. J., Saunderson, V., Black, S. & Davey, R. *J. Faraday Discuss.* **95**, 11–25 (1993).
- Boek, E. S., Briels, W. J., van Eerden, J. & Feil, D. *J. chem. Phys.* **96**, 7010–7018 (1992).
- Boek, E. S., Briels, W. J. & Feil, D. *J. phys. Chem.* **98**, 1674–1681 (1994).
- Boek, E. S., Briels, W. J. & Feil, D. *J. chem. Phys.* **98**, 1422–1428 (1993).
- Boek, E. S., Feil, D., Briels, W. J. & Bennema, P. *J. Crystal Growth* **114**, 389–410 (1991).
- Pickering, M. J. *Chem. Education* **64**, 723–724 (1987).

Chiral discrimination of monosaccharides using a fluorescent molecular sensor

Tony D. James,
K. R. A. Samankumara Sandanayake
& Selji Shinkai*

Chemirecognition Project, ERATO, Research Development Corporation of Japan, Aikawa 2432-3, Kurume, Fukuoka 830, Japan

MEANS of distinguishing between enantiomers of a chiral molecule are of critical importance in many areas of analytical chemistry and biotechnology, particularly in drug design and synthesis. In particular, solution-based sensor systems capable of chiral recognition would be of tremendous pharmaceutical value. Here we report the chiral discrimination of D- and L-monosaccharides using a designed receptor molecule that acts as a sensor by virtue of its fluorescent response to binding of the guest species. Our receptor contains boronic acid groups that bind saccharides by covalent interactions; such receptor systems have been much studied previously^{1–6} for complexation of saccharides, and have an advantage over others based on hydrogen-bonding interactions^{7–11}, for which polar protic solvents such as water can compete with guest binding. Our molecular sensor also incorporates a fluorescent naphthyl moiety; binding of each enantiomer of the monosaccharides alters the fluorescence intensity to differing degrees, enabling them to be distinguished. These water-soluble molecular sensors might form the basis of a quantitative and selective analytical method for saccharides.

Recently we developed a photoinduced-electron-transfer (PET)^{12,13} sensor for saccharides based on the interaction of boronic acid and amines^{2,3}. When saccharides form cyclic boronate esters with boronic acids such as **1** and **2** (Fig. 1), the acidity of the boronic acid is enhanced¹⁴ and therefore the Lewis acid–base interaction with the tertiary amine is strengthened. The strength of this acid–base interaction modulates the photoinduced

* To whom correspondence should be addressed.

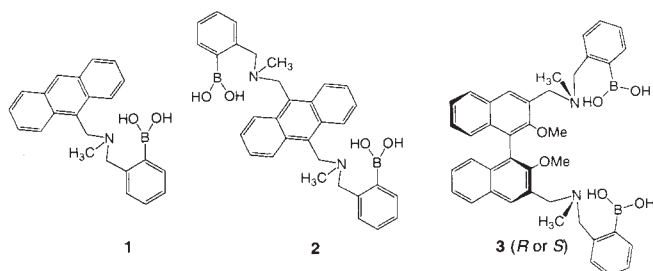


FIG. 1 Structures of photoinduced electron transfer molecular sensors.

ed-electron-transfer from the amine (acting as a quencher) to anthracene (acting as a fluorophore). Both these compounds show increased fluorescence at pH 7.7 through suppression of the photoinduced electron transfer from nitrogen to anthracene on saccharide binding; a direct result of the stronger boron-nitrogen bond^{2,3}.

The success in molecular design of this PET system prompted us to attempt to design a system capable of chiral recognition. From the seminal work by Irie^{15,16} on the control of intermolecular chiral 1,1'-binaphthyls fluorescence quenching by chiral amines, we realized that chiral recognition could involve both electronic and steric factors. We have used the 1,1'-binaphthyl moiety as a chiral building block. Although the synthesis of such an artificial receptor for saccharide discrimination might seem daunting, the recent synthesis of 2,2'-dimethoxy-1,1'-binaphthyl-3,3'-dialdehyde by Brunner¹⁷⁻¹⁹ offered a potential starting point in the synthesis of the chiral diboronic acid **3** (*R* or *S*). Brunner's 2,2'-dimethoxy-1,1'-binaphthyl-3,3'-dialdehyde¹⁷⁻¹⁹ was reductively aminated with methylamine followed by NaBH₄ to give the 3,3'-bis(*N*-methyl aminomethyl) derivative. This was then alkylated with 2-bromomethylphenylboronic acid (protected with 2,2-dimethyl-1,3-propanediol). The optical purity of the 2,2'-dimethoxy-1,1'-binaphthyl-3,3'-dialdehyde and **3** (*R* or *S*) were determined by high-performance liquid chromatography (HPLC). For the dialdehyde, we used a chirapak AD column with a mixed solvent of hexane and 2-propanol (4:1). For the diboronic acid we used a chirapak OP(+) column with methanol as solvent. The percentage enantiomeric excess (e.e.) and optical rotation for the *S* and *R* dialdehyde are 100% e.e. ($[\alpha]^{25}_D = +68.7^\circ$ ($c=0.14$, CHCl₃)) and 100% e.e. ($[\alpha]^{25}_D = -71.1^\circ$ ($c=0.4$, CHCl₃)), respectively. The percentage enantiomeric excess and optical rotation for the diboronic acid **3** (*R* or *S*) are >94% e.e. ($[\alpha]^{25}_D = +12.7^\circ$ ($c=0.272$, MeOH)) and >99% e.e. ($[\alpha]^{25}_D = -13.1^\circ$ ($c=0.4$, MeOH)), respectively. The HPLC peaks for the diboronic acid are broad and overlap slightly; lower limits for the enantiomeric excess were determined by addition of the other isomer.

TABLE 1 Stability constants and fluorescence enhancements for saccharides with **3R** (or *S*)

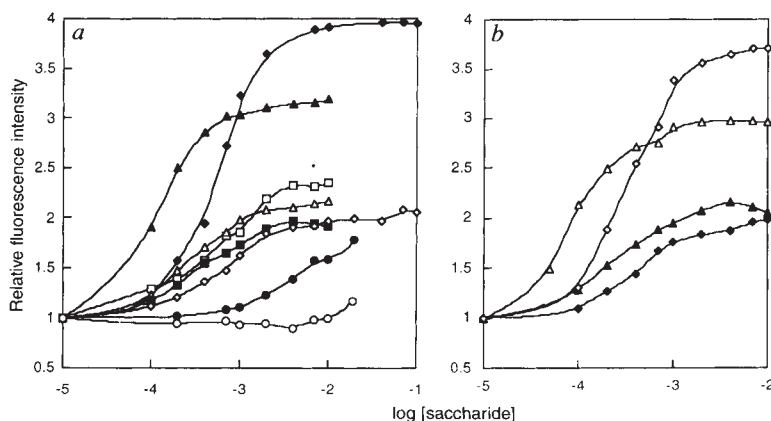
Saccharide	D log <i>K</i> (± 0.05)	L log <i>K</i> (± 0.05)	D/L fluorescence intensity ratio
Fructose	4.0 (3.7)	3.5 (4.0)	1.47 (0.69)
Glucose	3.3 (3.4)	3.1 (3.5)	1.93 (0.53)
Galactose	3.1	3.3	0.82
Mannose	<2.4	—	—

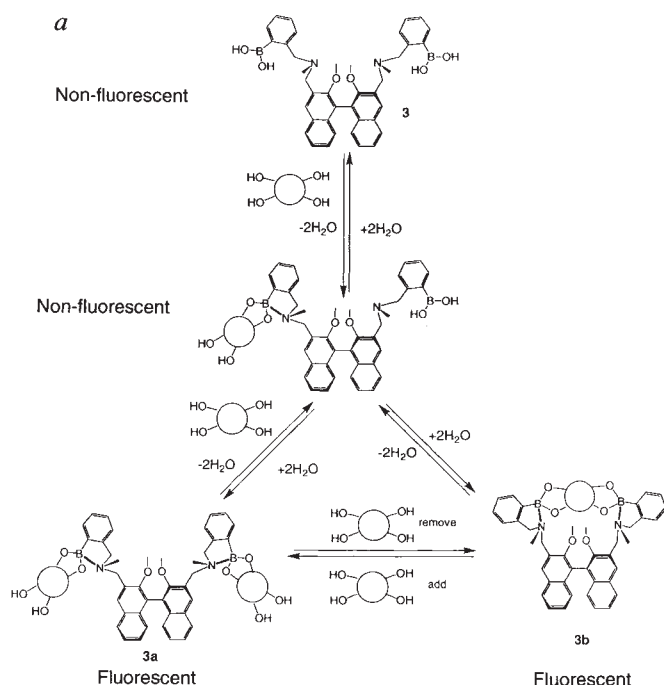
We hypothesized that, for **3** (*R* or *S*), chiral recognition might induce electronic effects. Electron transfer from the nitrogen to the naphthyl unit, and thus quenching efficiency, is dependent on their relative orientation. Formation of the 1:1 complex with a monosaccharide makes the boron-nitrogen interaction stronger and thus fixes the nitrogen in a certain orientation, which determines its quenching efficiency. Also, steric chiral recognition with **3** (*R* or *S*) is possible if an unfavourable interaction or twist around the binaphthyl bond occurs on binding D- or L-monosaccharides. Such a molecule might therefore be capable of two independent modes of chiral recognition.

The saccharide fluorescence intensity titration curves for **3** (*R* or *S*) at pH 7.77 (0.01000 M KCl, 0.002642 M KH₂PO₄, 0.00264 M Na₂HPO₄ 33.3% (w/w) methanol buffer)²⁰ are shown in Fig. 2. Figure 3a suggests which main species are involved. With **3** (*R* or *S*), four main species exist under the experimental conditions, and the fluorescent species are: **3a**, a non-cyclic 2:1 complex and **3b**, a cyclic 1:1 complex. Table 1 shows the stability constants, log *K*, for **3** (*R* or *S*) with monosaccharides. The stability constants were determined from the titration curves²¹ assuming the formation of a 1:1 complex. The stability constants determined for **3** (*R* or *S*) are in all cases greater than those determined with **1**, indicating that the two boronic acids are acting cooperatively in the formation of a 1:1 complex. Further evidence for the selective formation of a 1:1 complex is given by the mass (secondary ion mass spectroscopy positive) spectra of a 1:1 mixture of **3** (*R*) with D-glucose, D-mannose, D-fructose and D-galactose: all spectra contained the M⁺ ion of **3b** but the M⁺ ion of **3a** could not be detected.

Both steric and electronic chiral recognition are clearly demonstrated by **3** (*R* or *S*). With **3R** the 1:1 complex stabilities of D-fructose, D-glucose, D-mannose and L-galactose are greater than those of L-fructose, L-glucose, L-mannose and D-galactose, respectively (Table 1). With **3S** the 1:1 complex stabilities of L-fructose and L-glucose are greater than those of D-fructose and D-glucose, respectively (Table 1). We attribute these differences to steric effects. With **3R** the PET quenching efficiency on saccharide binding is affected most strongly by D-fructose, D-glucose and L-galactose; the least by L-fructose, L-glucose and

FIG. 2 a, Fluorescence intensity log [saccharide] profile of **3R** at 25 °C; 1.0×10^{-5} M of **3R** in 33.3% MeOH, H₂O buffer at pH 7.77, excitation wavelength at 289 nm, emission wavelength at 358 nm (λ_{ex} 289 nm, λ_{em} 358 nm). b, Fluorescence intensity log [saccharide] profile of **3S** at 25 °C; 1.0×10^{-5} M of **3S** in 33.3% MeOH, H₂O buffer at pH 7.77, λ_{ex} 289 nm, λ_{em} 358 nm. D-Glucose (◆); L-glucose (◇); D-galactose (■); L-galactose (□); D-fructose (▲); L-fructose (△); D-mannose (●); L-mannose (○).





D-galactose (Table 1). With **3S** the PET quenching efficiency on saccharide binding is affected most strongly by L-fructose and L-glucose; the least by D-fructose and D-glucose (Table 1). This demonstrates the electronic effects of complexation.

Previously, we demonstrated (with **2**) that the monosaccharide selectivity of a diboric acid could be finely tuned towards one particular monosaccharide (glucose). With **3** we have now shown that selectivity towards one optical isomer of a monosaccharide can also be achieved. Furthermore, we are able to demonstrate selective recognition of one enantiomer in the presence of the other (Fig. 3b). We propose that the PET sensor **3** is a 'bimodal' chiral probe: both the fluorescence intensity and the stability of the 1:1 saccharide complex reflect the chirality of the bound saccharide. □

Received 25 November 1994; accepted 10 February 1995.

- Deng, G., James, T. D. & Shinkai, S. *J. Am. chem. Soc.* **116**, 4567–4572 (1994).
- James, T. D., Sandanayake, K. R. A. S. & Shinkai, S. *J. chem. Soc., chem. Commun.* 477–478 (1994).
- James, T. D., Sandanayake, K. R. A. S. & Shinkai, S. *Angew. Chem.* **33**, 2207–2209 (1994).
- Yoon, J. & Czarnik, A. W. *J. Am. chem. Soc.* **114**, 5874–5875 (1992).
- Morin, G. T., Hughes, M. P., Paugam, M.-F. & Smith, B. D. *J. Am. chem. Soc.* **116**, 8895–8901 (1994).
- Wulff, G. *Pure appl. Chem.* **54**, 2093–2102 (1982).
- Rebek, J. Jr *Angew. Chem.* **29**, 245–255 (1990).
- Goswami, S. & Hamilton, A. D. *J. Am. chem. Soc.* **111**, 3425–3426 (1989).
- Kelly, T. R., Zhao, C. & Bridger, G. J. *J. Am. chem. Soc.* **111**, 3744–3745 (1989).
- Etter, M. C. & Admond, D. A. *J. chem. Soc., chem. Commun.* 589–591 (1990).
- Aoyama, Y., Tanaka, Y., Toi, H. & Ogoshi, H. *J. Am. chem. Soc.* **110**, 634–635 (1988).
- Bissel, R. A. et al. *Top. Curr. Chem.* **168**, 223–264 (1993).
- Czarnik, A. W. *Fluorescent Chemosensors for Ion and Molecular Recognition* (ACS Books, Washington DC, 1993).
- Lorand, J. P. & Edwards, J. D. *J. org. Chem.* **24**, 769–774 (1959).
- Irie, M., Yorozu, T. & Hayashi, K. *J. Am. chem. Soc.* **100**, 2236–2237 (1978).
- Yorozu, T., Hayashi, K. & Irie, M. *J. Am. chem. Soc.* **103**, 5480–5484 (1981).
- Brunner, H. & Goldbrunner, J. *Chem. Ber.* **122**, 2005–2009 (1989).
- Brunner, H. & Wulz, K. *New J. Chem.* **16**, 57–61 (1992).
- Brunner, H. & Schiessling, H. *Angew. Chem.* **33**, 125–126 (1994).
- Perrin, D. D. & Dempsey, B. *Buffers for pH and Metal Ion Control* Ch. 6 (Chapman & Hall, London, 1974).
- Fery-forgues, S., LeBris, M.-T., Guette, J.-P. & Valeur, B. *J. phys. Chem.* **92**, 6233–6237 (1988).

ACKNOWLEDGEMENTS. We thank M. Irie (Kyushu University) for helpful discussions; T.D.J. and K.R.A.S.S. thank P. Linnane and M. Mikami (JRC Chemecognics Project) for helpful advice; T.D.J. thanks H. Brunner for supplying detailed experimental procedures. We also thank R. Iguchi and S. Imazu for technical assistance.

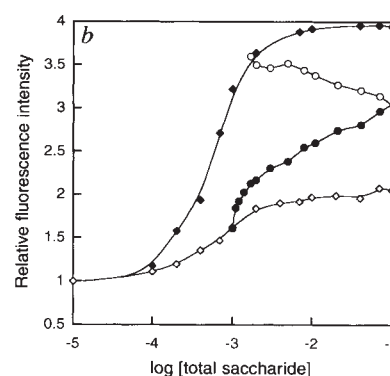


FIG. 3 **a**, Possible complexes of **3** (*R* or *S*) with saccharides at pH 7.77. **b**, Chiral recognition for mixtures of enantiomers. The figure shows the fluorescence intensity log [total saccharide] profile of **3R** at 25 °C; 1.0×10^{-5} M of **3R** in 33.3% MeOH/H₂O buffer at pH 7.77, λ_{ex} 289 nm, λ_{em} 358 nm. D-Glucose (◆); L-glucose (◇); added D-glucose from initial condition of 0.001 M L-glucose (●); added L-glucose from initial condition of 0.001 M D-glucose (○).

Palaeoaltimetry from energy conservation principles

Chris E. Forest, Peter Molnar & Kerry A. Emanuel

Center for Meteorology and Physical Oceanography,
Massachusetts Institute of Technology, Cambridge,
Massachusetts 02139, USA

A KNOWLEDGE of past changes in the mean elevations of large continental areas is important for understanding both dynamic processes in the Earth's mantle¹ and the evolution of the global climate². But virtually all methods for determining palaeoelevations are problematic³, in part because changes in either elevation or climate can give rise to the same observed geological phenomena⁴. Ideally, palaeoelevations would be inferred directly from estimates of palaeopressure, and it has recently been shown⁵ that vesicles in basaltic lava flows preserve a record of atmospheric pressure at the altitude of emplacement; however, the elevations thus obtained have large errors (~1.4 km), and the method can at present be applied only to lavas that have had a simple emplacement history⁵. Palaeobotanical methods for estimating palaeoelevation have received much attention^{6–12}, but they rely on empirical temperature–elevation relationships, the uncertainties in which are difficult to evaluate for past climates. Here we describe an alternative palaeobotanical approach, based on energy conservation in the atmosphere, in which fossil leaf assemblages are used to infer enthalpy, rather than temperature. This approach is relatively insensitive to palaeoclimate, with an expected error in palaeoelevation of ~700 m.

The use of fossil leaf assemblages for palaeoaltimetry consists of two steps: first, estimating from fossil plants a climate parameter that varies with altitude (like mean annual temperature), and second, using differences in that parameter from separate sites to estimate elevation differences⁶, given *a priori* knowledge of the variation of the parameter with altitude. Mean annual temperature has been used predominantly as the climate param-

eter because it shows correlations with physiognomic characteristics of leaf samples from living plants^{13–16}. Assuming that characteristics of fossil leaf assemblages obey the same relationship, palaeotemperature estimates can be inferred^{7,17} with uncertainties perhaps as low as 1 °C. As mean annual temperature varies with altitude, latitude and geological time, palaeoelevation (Z) may be calculated by comparing surface temperatures, T , of high- and low-altitude locations at the same latitude and of the same age using:

$$Z = (T_{\text{high}} - T_{\text{low}}) \gamma^{-1} \quad (1)$$

where $\gamma = -\partial T / \partial Z$ is the empirical relationship between mean annual temperature at the surface and altitude derived from the present climate. (Note that γ is not a gradient of T within the atmosphere and so is not a lapse rate in the meteorological sense, although it has been termed a 'terrestrial lapse rate'.)

Not only do relationships between T and Z lack a firm theoretical foundation, but large spatial variations also exist in present-day estimates of γ in the western United States^{8–10}. Meyer¹⁰ calculated local estimates of γ for 39 areas of the world with surface topography >750 m and found $\gamma = 5.9 \pm 1.1 \text{ K km}^{-1}$, but with a range of 3.64–8.11 K km^{-1} . Moreover, we have no reason to expect that the present laterally varying empirical relationships should hold in the different climates that occurred in geological time. We therefore seek a measurable thermodynamic quantity whose distribution with altitude and longitude is constrained well by both theory and observations. The use of a quantity chosen to agree with fundamental thermodynamic laws is obviously more reliable than one fitted empirically to data spanning a fraction of the twentieth century—a small fraction of geological history.

Derived from the first law of thermodynamics, moist static energy^{18,19}, h , is the total specific energy content of air ($\sim 300 \text{ kJ kg}^{-1}$ in mid-latitudes), excluding kinetic energy, which is very small (<5%) compared with the other terms:

$$h = c_p' T + L_v q + gZ = H + gZ \quad (2)$$

where c_p' is the specific heat capacity at constant pressure of moist air, T is temperature (in K), L_v is the latent heat of vaporization for water, q is specific humidity, g is the gravitational acceleration, Z is altitude and H is moist enthalpy. The specific heat capacity of moist air, $c_p' = (1-q)c_{pd} + qc_{pv}$, accounts for compositional changes of the air, where c_{pd} and c_{pv} are the speci-

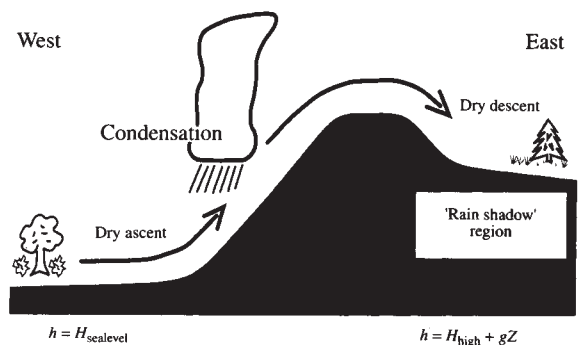
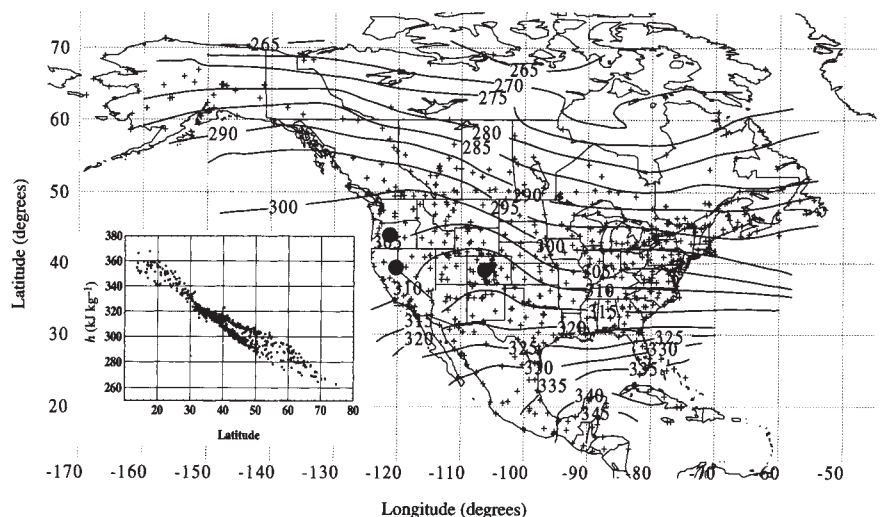


FIG. 1 Surface air conserves moist static energy as it traverses a mountainous region through conversions between sensible heat, latent heat and potential energy. During dry ascent and descent, changes in sensible heat compensate for changes in potential energy. For condensation processes at a constant height, the sensible heat increases when the latent heat of water vapour is released. The potential energy for the high-elevation site is the difference in moist enthalpy, $\Delta H = H_{\text{sealevel}} - H_{\text{high}}$, which, via $\Delta H/g$, yields the elevation estimate.

fic heat capacities of dry air and water vapour respectively. Although the internal energy, $c_p' T$, is the largest term in equation (2), typical variations in the three separate components are roughly of equal magnitudes, ~ 30 – 50 kJ kg^{-1} , for the ranges of T , q and Z of approximately 0–30 °C, 0–20 g kg^{-1} , and 0–4,000 m respectively.

Two properties of moist static energy make it a desirable candidate for inferring palaeoelevations. First, changed only by radiative heating and surface fluxes of latent and sensible heat, it is nearly conserved following air parcels and hence is approximately constant along trajectories. Second, the value of h in the atmospheric boundary layer (the lowest 1–1.5 km) is usually constrained by convection to be close to the value of h in the upper troposphere^{20,21}. (The actual constraint is along vortex lines rather than in the vertical direction, but the difference is usually small.) The Earth's rotation causes large-scale movement in the upper troposphere to flow nearly from west to east at mid-latitudes, so that contours of h there should also trend nearly west–east. This implies that h at the surface should be nearly invariant with longitude (that is, be zonally symmetric).

FIG. 2 The spatial distribution of mean annual moist static energy. Crosses indicate station locations from the ISMCS dataset²². Mean moist static energy h was calculated by independently calculating a mean temperature and mean specific humidity and by combining the respective sensible heat, latent heat and potential energies. We accounted for the temperature dependence of the latent heat of vaporization, $L_v = L_{v0} + (T - 273.15)(c_{pv} - c_l)$, where c_{pv} and c_l are the specific heat capacity at constant pressure of water vapour and liquid water, respectively, and L_{v0} is the latent heat of vaporization at 0 °C. This second-order effect and the dependence of c_p' on q are most important in warm and moist climates, where they contribute variations of ~ 5 – 10 kJ kg^{-1} corresponding to apparent differences in elevation of 500–1,000 m. Of the various averaging periods, we chose mean annual h based on the small error from zonal asymmetries and on the ability to predict mean annual enthalpy from fossil leaf physiognomy. The locations of the fossil collection sites are indicated by filled circles: Goshen, Oregon; LaPorte, California; and Florissant, Colorado. Inset, the variation of h with latitude (in degrees). Limiting the data to 30°–60° N yields the estimate of $s_h = 4.5 \text{ kJ kg}^{-1}$.



As a caveat, the large-scale atmospheric flow is influenced by topography that will affect the zonal symmetry. So when applying this technique, one should ascertain the zonal symmetry of the palaeoclimate; this requires a better understanding of stationary waves in the general circulation of the atmosphere.

Assuming that h is zonally symmetric along the Earth's surface, if we can estimate enthalpy at sea level ($H_{\text{sea level}}$) for a particular latitude, it follows from equation (2) that the palaeo-altitude, Z , of another location at the same latitude is given by

$$Z = \frac{H_{\text{sea level}} - H_{\text{high}}}{g} \quad (3)$$

where H_{high} is the enthalpy at the high-altitude locations (Fig. 1). Therefore, the necessary assumptions to estimate palaeo-altitude are: (1) the surface moist static energy is zonally symmetric, and (2) the estimate of enthalpy for the regional palaeoclimate is accurate.

To examine deviations from zonal invariance of h , we examined the distribution of mean h across the North American continent (Fig. 2). Using climate data for 699 surface weather stations²², we calculated mean h by independently calculating a mean temperature and mean specific humidity and by combining the respective sensible heat, latent heat and potential energies. We used averaging periods ranging from a season to a year, and various formulations of mean temperature and mean specific humidity. We quantified the longitudinal invariance by statistically fitting h to monotonic arbitrary functions of latitude, ϕ . Because we are testing for invariance, the standard deviation of the fit is more important than the actual latitude dependence. To minimize the expected error, we limited the data to 30°–60° N and divided them into three latitude bins of 10°. We linearly fitted the data within each band and used deviations from each to obtain a mean square error (Fig. 2). This approach reduces the expected error from zonal asymmetries of h , s_h , to 4.5 kJ kg⁻¹. Dividing s_h by g yields an estimate of the minimum error in altitude of 460 m.

We can compare this error to that deduced assuming the invariance of the terrestrial lapse rate, and therefore the longitudinal invariance of $G = T + \gamma Z$. For $\gamma = 5.9 \text{ K km}^{-1}$ (ref. 10), we calculated deviations from longitudinal invariance of G to be 3.2 K for the present-day climate of North America. Hence a minimum

standard error for estimating altitude can be computed from these results, from $s_Z^2 = s_G^2/\gamma^2 + ((\Delta T)^2 s_\gamma^2)/\gamma^4$ where s_G and s_γ are the standard errors of G and γ respectively. With $s_\gamma = 1.1 \text{ K km}^{-1}$ (ref. 10), minimum uncertainties in the estimated altitudes increase with altitude: $s_Z = 540 \text{ m}$, 660 m and 920 m for altitudes of 0 m, 2,000 m and 4,000 m, respectively. So even ignoring the unpredictable temporal changes in γ over geological time, these estimates of minimum uncertainties exceed that of 460 m from assuming longitudinal invariance in h .

The task remains to determine mean annual enthalpy from plant physiognomy. We present results of an analysis relating foliar physiognomic characters to mean annual enthalpy that exploits the method and data developed by Wolfe⁷. We searched among an array of 29 foliar characteristic scores (p_i ; see below) from 92 sites in North America for linear combinations of the foliar characteristics that covary with the local climates. At a given site, the scores are calculated as a percentage of species exhibiting each of 29 foliar characteristics: seven ranges of size, five length-to-width ratios, overall shape (obovate, elliptic, ovate, lobed), basal shape (cordate, round, acute), apical shape (emarginate, rounded, acute, attenuate), and teeth (no teeth, regular, close, round, acute, compound). We determined which foliar characteristics covary with one another (using principal components analysis) and which best correlate with enthalpy (using principal components analysis with regression²³). This creates a predictive equation of the form:

$$H = H_0 + \sum_i a_i p_i \quad (4)$$

where a_i is the regression coefficient of the i th foliar characteristic. The standard error for predicting H is 3.9 kJ kg⁻¹ (Fig. 3), corresponding to a minimum error in palaeoelevation of 400 m.

To illustrate the use of this method in determining palaeoelevation, we estimated H from the physiognomic characteristics for three well known fossil floral assemblages: Florissant (34.1 Myr) in central Colorado, Goshen (32 Myr) from central Oregon and LaPorte (33 Myr) from northern California (Fig. 2). Each assemblage of fossil leaves has been appropriately scored using the same characteristics as in Wolfe's data set⁷ and has been used in previous estimates of Florissant's palaeoelevation including 2.3–3.3 km (refs 11, 12), 2.5 km (ref. 8) and 2.9 km (ref. 9). Using Wolfe's scores (personal communication) for the Goshen and LaPorte floras and Gregory's (ref. 11) for the Florissant flora, we calculated mean enthalpies of 294.6, 321.1 and 321.2 kJ kg⁻¹, respectively, for these assemblages. Assuming that both Goshen and LaPorte were deposited at sea level yields a sea-level value of mean enthalpy of $321.1 \pm 2.8 \text{ kJ kg}^{-1}$. The difference in enthalpy between sea level and Florissant yields a palaeoelevation of $2,700 \pm 670 \text{ m}$. Sea level 32–34 Myr ago was 180 m higher²⁴ than at present, so the palaeoelevation is $2.9 \pm 0.7 \text{ km}$ with respect to the present geoid. This estimate of Florissant's palaeoaltitude agrees with other recent estimates of palaeoaltitude, and therefore supports the contention of Gregory and Chase^{11,12}, Meyer⁸ and Wolfe⁹, that Florissant was high in Late Eocene/Early Oligocene time. As these workers have noted, there is a widespread belief that the Rocky Mountains were at a low elevation throughout much of the Late Cenozoic and rose early in Quaternary times. More importantly here, the 700-m uncertainty is smaller than those of other palaeoelevation estimates. □

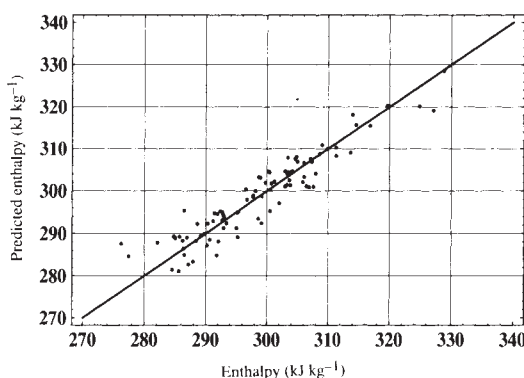


FIG. 3 The predicted mean enthalpy versus observed mean enthalpy for the annual mean, as given by equation (4) for the 92 vegetation sites (see text). The regression analysis accounts for 88% of the variations in moist enthalpy, yielding a standard error $s_H = 3.9 \text{ kJ kg}^{-1}$. Because humidity data are not available for each vegetation site, mean annual enthalpy for the plant collection sites is interpolated from nearby stations using the standard distance-weighted mean, $h_{\text{plant site}} = (\sum w_i h_i) / (\sum w_i)$ where $w_i = \exp(-r_i^2/d^2)$, where r_i is the radial distance from plant site to the meteorological station i , and d is the average distance to the nearest station, 90.8 km. Knowledge of the elevation, Z , and mean annual temperature, T , at the vegetation sites allows the calculation of mean moist enthalpy, $H = h - gZ$, from the interpolated value of h .

Received 7 November 1994; accepted 17 February 1995.

- England, P. & Houseman, G. J. *geophys. Res.* **94**, 17561–17579 (1989).
- Raymo, M. E. & Ruddiman, W. F. *Nature* **359**, 117–122 (1992).
- England, P. & Molnar, P. *Geology* **18**, 1173–1177 (1990).
- Molnar, P. & England, P. *Nature* **346**, 29–34 (1990).
- Sahagian, D. L. & Maus, J. E. *Nature* **372**, 449–451 (1994).
- Axelrod, D. I. *Paleobotanist* **14**, 144–171 (1966).
- Wolfe, J. A. *Bull. U.S. geol. Surv.* **2040**, (1993).
- Meyer, H. W. thesis, Univ. California at Berkeley (1986).
- Wolfe, J. A. *Bull. U.S. geol. Surv.* **1964** (1992).
- Meyer, H. W. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **99**, 71–99 (1992).

11. Gregory, K. M. *Paleoclimates* (in the press).
12. Gregory, K. M. & Chase, C. G. *Geology* **20**, 581–585 (1992).
13. Bailey, I. W. & Sinnot, E. W. *Science* **41**, 831–834 (1915).
14. Bailey, I. W. & Sinnot, E. W. *Am. J. Bot.* **3**, 24–39 (1916).
15. Wolfe, J. A. & Hopkins, H. E. in *Tertiary Correlation and Climatic Changes in the Pacific* (ed. Hatai, K.) 67–76 (Sasaki, Sendai, 1967).
16. Wolfe, J. A. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **9**, 27–57 (1971).
17. Wolfe, J. A. *Prof. pap. U.S. geol. Surv.* **1106** (1979).
18. Wallace, J. M. & Hobbs, P. V. *Atmospheric Science: An Introductory Survey* (Academic, New York, 1977).
19. Emanuel, K. A. *Atmospheric Convection* (Oxford Univ. Press, New York, 1994).
20. Betts, A. K. J. *atmos. Sci.* **39**, 1484–1505 (1982).
21. Xu, K.-M. & Emanuel, K. A. *Mon. Weath. Rev.* **117**, 1471–1479 (1989).
22. International Station Meteorological Climate Summary CD-ROM Version 1.0 (Federal Climate Complex, Asheville, NC, 1990).
23. Mardia, K. V., Kent, J. T. & Bibby, J. M. *Multivariate Analysis* (Academic, New York, 1979).
24. Haq, B. U., Hardenbol, J. & Vall, P. R. *Science* **235**, 1156–1167 (1987).

ACKNOWLEDGEMENTS. We thank J. Wolfe for his contributions of data and expert advice, which made the entire plant physiognomic/climate correlations possible. We thank K. Gregory for unpublished Florissant data, A. Del Genio and K. Gregory for reviews, M. Raymo for revisions of early drafts, and R. Spicer, P. England, D. Povey and L. Stranks for discussions about palaeobotany and multivariate analysis. This work was supported by the Climate Dynamics Program in the Division of Atmospheric Sciences of the US NSF and in part by an ONR fellowship (C.E.F.).

Dorsalizing signal *Wnt-7a* required for normal polarity of D–V and A–P axes of mouse limb

Brian A. Parr & Andrew P. McMahon

Department of Molecular and Cellular Biology, Harvard University, 16 Divinity Avenue, Cambridge, Massachusetts 02138, USA

FORMATION of the vertebrate limb requires specification of cell position along three axes¹. Proximal–distal identity is regulated by the apical ectodermal ridge (AER) at the distal tip of the growing limb^{2–6}. Anterior–posterior identity is controlled by signals from the zone of polarizing activity (ZPA) within the posterior limb mesenchyme^{7–9}. Dorsal–ventral identity is regulated by ectodermally derived signals^{10–14}. Recent studies have begun to identify signalling molecules that may mediate these patterning activities. Members of the fibroblast growth factor (FGF) family are expressed in the AER and can mimic its proximal–distal signalling activity^{15,16}. Similarly, the gene *Sonic hedgehog* (*Shh*) is expressed in the ZPA, and *Shh*-expressing cells, like ZPA cells, can cause digit duplications when transplanted to the anterior limb margin^{17,18}. In contrast, no signal has yet been identified for the dorsal–ventral axis, although *Wnt-7a* is expressed in the dorsal ectoderm, suggesting that it may play such a role^{19,20}. To test this possibility, we have generated mice lacking *Wnt-7a* activity. The limb mesoderm of these mice shows dorsal-to-ventral transformations of cell fate, indicating that *Wnt-7a* is a dorsalizing signal. Many mutant mice also lack posterior digits, demonstrating that *Wnt-7a* is also required for anterior–posterior patterning. We propose that normal limb development requires interactions between the signalling systems for these two axes.

Wnt-7a is expressed in the flanking ectoderm of the trunk prior to limb bud outgrowth. At 8.75 days postcoitum (d.p.c.), *Wnt-7a* expression encompasses the presumptive forelimb region, extending from the level of the last few somites into the ectoderm overlying the presumptive mesoderm (Fig. 1a). At the level of the presumptive hindlimb, ectodermal expression is first apparent at 9.25 d.p.c. (Fig. 1a). During initial stages of limb-bud outgrowth (9.25 d.p.c., forelimb; 9.75 d.p.c., hindlimb) and patterning, *Wnt-7a* transcripts are uniformly distributed throughout the dorsal limb ectoderm (Fig. 1a).

To investigate the function of *Wnt-7a*, we used gene targeting in mouse embryonic stem (ES) cells to generate a likely null allele (Fig. 1b–d). Matings between mice heterozygous for the targeted allele produced homozygous offspring at expected men-

delian frequencies (+/+, 32; +/-, 70; -/-, 37). Although apparently fully viable, *Wnt-7a*^{-/-} mice have limb abnormalities and are also sterile.

Ventral structures develop normally in *Wnt-7a*^{-/-} limbs, but many dorsal tissues adopt ventral fates. Distal limb structures are particularly useful in distinguishing dorsal from ventral cell types (Figs 2, 3). The ventral skin of wild-type digits is devoid of fur, lacks pigmentation and exhibits a prominent set of transverse striations (Fig. 2a). Moreover, dermal footpads protrude at the base of the digits and at the distal tips of the digits (Fig. 2a). Normal flexion of the digits depends upon a specific arrangement of ventral tendons projecting to a series of small, ventrally located, sesamoid bones. In contrast, the dorsal half of the limb is pigmented, covered in hair, and does not possess footpads or sesamoid bones (Fig. 2b). The major dorsal tendons travel along the outside edges of the digits (Fig. 3b), while a larger tendon runs down the middle of the ventral surface (Fig. 3a). Dorsal-to-ventral transformations of mesenchyme cell fate are observed in the footpads, skin, tendons, sesamoid bones and joints of *Wnt-7a* mutant limbs.

The most striking superficial abnormalities on the dorsal surface of mutant limbs are dermal thickenings representing duplications of the footpads normally found ventrally (Fig. 2c, e, f). The dorsal pads are visible as early as 15.5 d.p.c. (Figs 2g–i, 3d, e) and express *Pax-9* (Fig. 2i), a putative transcription factor, normally restricted to developing ventral pads (A. Neubuser and R. Balling, personal communication). Ectopic dorsal pads are pigmented, as migration of melanocytes is apparently unaffected by the *Wnt-7a* mutation.

The dorsal surfaces of the digits in severe *Wnt-7a* mutants also lose hair and acquire noticeable striations, resembling the normal ventral surface (Fig. 2c). The growth of nails at the dorsal–ventral interface of the digits is truncated to a variable extent (Fig. 2c, f), and abnormal, pigmented thickenings, which probably represent duplications of the large ventral pads found at the distal ends of the digits (Fig. 2a), often grow over the rudimentary nails (Fig. 2f).

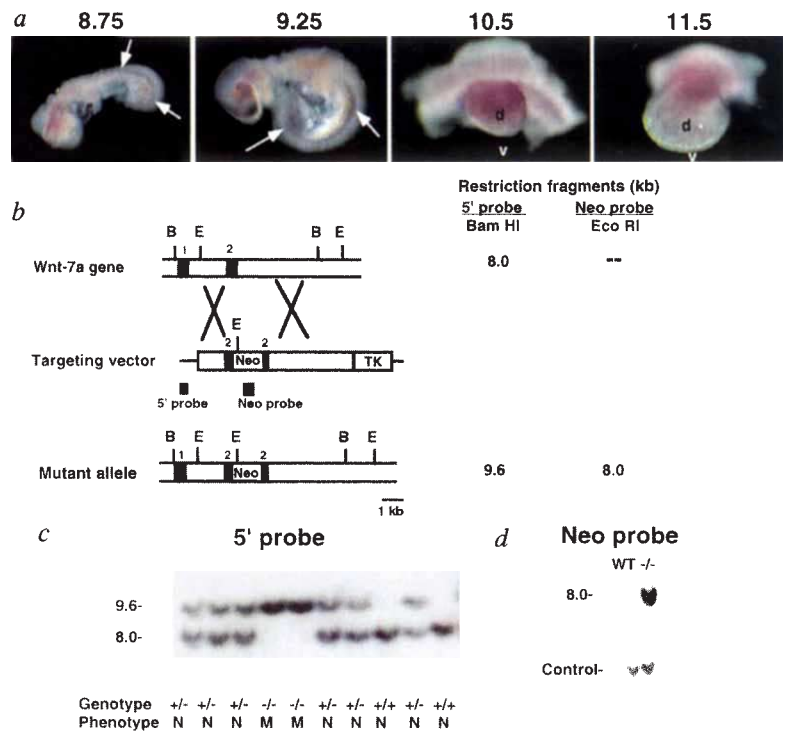
Alterations in dorsal–ventral polarity are also apparent in the tendons and bones. In adult *Wnt-7a* mutant limbs (Fig. 3c), a large dorsal tendon, closely resembling the central ventral tendons (Fig. 3a), projects distally along each digit. These duplications of ventral tendons are clearly visible in sections through 15.5 d.p.c. mutant limbs, where the dorsal tendons appear to be mirror images of ventral tendons (compare Fig. 3d, f with e, g).

Normal flexion of the digits requires the correct arrangement of tendons and their attachment to sesamoid processes on the ventral surface of the distal limb bones. Sesamoid bones appear in their normal ventral positions in *Wnt-7a* mutants (Fig. 3i), but paired sesamoid processes also develop in the dorsal half of mutant paws (Fig. 3k, n, o). Not surprisingly, *Wnt-7a* mutant limbs show abnormal flexion due to these alterations in dorsal–ventral polarity. In contrast to the flexion of wild-type digits (Fig. 3l), mutant digits are generally straight (Fig. 3n), presumably because the duplication of ventral tendons dorsally results in similar tendons exerting counterbalancing forces on either side of the joints. In addition, the paws of mutants usually splay downwards and outwards, indicating incorrect bending at both the wrist and elbow joints (Fig. 3p, q).

Although we consistently observe a ventralization of the dorsal limb, the phenotype is clearly more severe distally than proximally. Interestingly, experiments in the chick indicate that ectodermal signalling only regulates dorsal–ventral patterning in the distal limb^{13,14}. Thus, whereas *Wnt-7a* is important in dorsal–ventral patterning, it is likely that other mechanisms are also involved.

To determine whether changes in ectodermal cell fate contribute to the observed phenotype, we examined the expression of a number of regionally restricted ectodermal markers between 9.5 and 10.5 d.p.c. *Wnt-7a* transcripts can be detected in the dorsal ectoderm of *Wnt-7a*^{-/-} limbs (data not shown), sug-

FIG. 1 a, Whole-mount *in situ* hybridization analysis of *Wnt-7a* expression from 8.75 d.p.c. to 11.5 d.p.c. *Wnt-7a* is expressed in the flank ectoderm (between the arrows) overlying the last few somites and extending into the presomitic region at 8.75 d.p.c., in the ectoderm overlying the presumptive limb bud regions at 9.25 d.p.c. (arrows), and uniformly throughout the dorsal ectoderm at 10.5–11.5 d.p.c. (dorsal views; d, dorsal; v, ventral). **b**, Targeting strategy to mutate the *Wnt-7a* gene by inserting a neomycin-resistance (*neo*) gene into the second exon. **c, d**, Southern blots demonstrating correct targeting of the *Wnt-7a* locus and genotyping of progeny (N, normal; M, mutant) using 5' and *neo* probes indicated in **b**. **METHODS**. Whole-mount *in situ* hybridizations were done as previously described¹⁹. *Wnt-7a* genomic clones were isolated from a 129/Sv library (Stratagene) using a cDNA probe spanning the entire coding region. A double-selection gene-replacement construct was designed to insert a *neo* gene into an *NaeI* site in the second exon of the *Wnt-7a* gene. *Wnt-7a* transcripts will be detected in mutant mice carrying this allele. The construct was linearized with *Sall* and electroporated into CJ7 ES cells³¹. 96 G418/FIAU resistant colonies were picked and screened by Southern blot analysis. Filters hybridized with the *neo* probe were subsequently reprobated with a control *Wnt-7a* fragment that recognized both the wild-type and targeted alleles. Two independent targeting events were identified, and the targeted clones were injected into C57B1/6J blastocysts to generate chimaeric mice. Tail or yolk sac DNA samples from offspring were initially genotyped by the same Southern protocol used to identify targeted clones. A PCR strategy was later devised to distinguish between targeted and wild-type alleles.



gesting that the dorsal ectoderm itself is not ventralized. Similarly, expression of *En-1*, *Bmp-2*, and *Dlx-2*, normally restricted to the ventral ectoderm^{21–23}, remains ventral in mutant limb buds (data not shown). Thus, cell fate changes would appear to be restricted to the underlying dorsal mesenchyme. These results also suggest that if a second ectodermal signal is required to specify ventral mesenchyme development, it is already present in the dorsal ectoderm, but the ventralizing activity is normally blocked at some regulatory level by *Wnt-7a*. Such a ventralizing signal, presumably acting on cells in the progress zone, may be widely expressed in the ectoderm or restricted to the dorsal-ventral interface in the region of the AER.

Wnt-7a also appears to regulate patterning along the anterior-posterior axis. There is a notable loss of posterior skeletal elements in mutant mice (Fig. 4a), even though *Wnt-7a* expression appears to be uniform throughout the dorsal ectoderm. Digit five, the most posterior digit, and the ulna, the more posterior zeugopod element of the forearm, are often missing or abnormal in *Wnt-7a* homozygotes. The frequency of digit abnormalities increases in a graded fashion from anterior to posterior (Fig. 4b), as does the severity of defects within any individual limb. The extent of posterior abnormalities is quite variable (Fig. 4b), although part of the variability may result from the non-uniform genetic background of the animals. Loss of posterior bones is

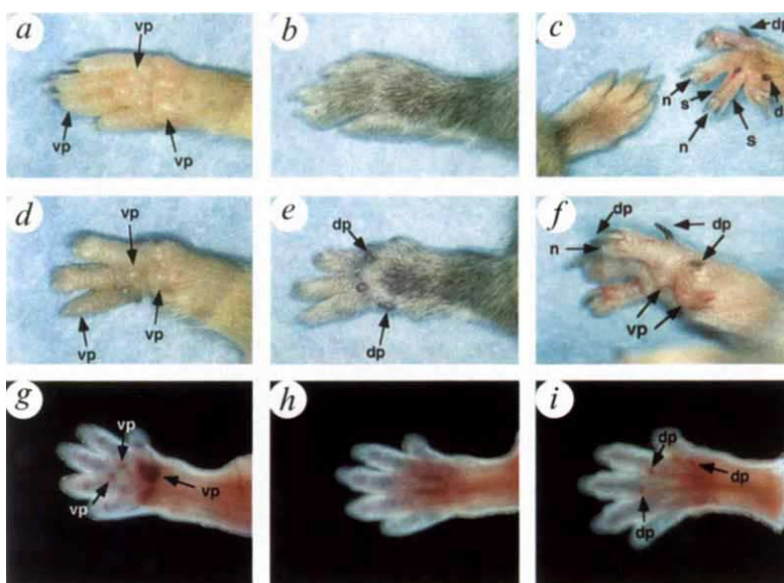


FIG. 2 Superficial views of *Wnt-7a* mutant limbs demonstrating the duplication of dermal pads on the dorsal surface of the paws. **a, d**, Ventral surface of wild-type (**a**) and mutant (**d**) forelimbs two weeks after birth showing the normal ventral pads (vp). **b, e**, Dorsal surface of wild-type (**b**) and mutant (**e**) forelimbs two weeks after birth showing ectopic dorsal pads (dp) on the mutant limbs. **c**, Wild-type and mutant adult forelimbs. Note the duplicated dorsal footpads (dp), the striations and lack of hair on the skin (s), and the abnormal nails (n) of the mutant limb. **f**, Close-up of adult mutant forelimb. The ectopic dorsal footpads (dp) maintain the same anterior-posterior and proximal-distal positions as the normal ventral pads (vp). The mutant nails (n) are truncated and overgrown by a duplication of the distal footpads. **g–i**, Whole-mount *in situ* hybridization demonstrating *Pax-9* expression in developing pads of 15.5-d.p.c. embryos. **g**, Wild-type ventral pads (vp) express *Pax-9*. **h**, No *Pax-9* expression is apparent on the dorsal surface of wild-type limbs. **i**, The dorsal surface of *Wnt-7a* homozygous limb buds shows ectopic *Pax-9* expression in developing dorsal pads (dp).

not accompanied by any apparent transformation of posterior structures towards an anterior phenotype.

The loss of posterior skeletal elements in mutant limbs suggests that *Wnt-7a* signalling may be necessary for normal ZPA function. Attenuation of ZPA activity, either by grafting fractions of the ZPA or irradiation of normal ZPA cells, also leads to the preferential loss of posterior skeletal structures^{9,24}. One potential target of *Wnt-7a* activity is *Shh*, which is expressed in the posterior mesenchyme and appears to mediate ZPA function^{18,25}. *Wnt-7a* expression in the forelimb bud precedes *Shh* expression, which is not visible until 9.75 d.p.c.¹⁷. Wild-type *Shh* expression in the posterior mesenchyme is strongest in the dorsal half of the limb bud (Fig. 4c), suggesting that a dorsally localized factor may regulate activity of the ZPA. Initial expression of *Shh* in the forelimb bud is detected at the normal time in *Wnt-7a* mutants (not shown), so that activation of the *Shh* gene does not depend upon *Wnt-7a* function. However, by 10.5 d.p.c. there are fewer *Shh*-expressing cells in the limb buds of most mutant embryos (Fig. 4d, e). The decrease in *Shh* expression at a given stage varies from embryo to embryo, resembling

the variable loss of posterior skeletal elements observed in older animals and suggesting that the levels of *Shh* activity may regulate anterior-posterior polarity. Thus, animals with the lowest levels of *Shh* expression in the early limb bud may subsequently develop limbs with the most severe posterior truncations. BMP-2, a member of the transforming growth factor TGF- β family, is also expressed in the posterior mesenchyme at 10.5 d.p.c.²¹. Like *Shh*, *Bmp-2* expression in the ZPA is mainly dorsal and is also lost to a variable extent in *Wnt-7a* mutants (data not shown). Thus, *Wnt-7a* activity is necessary for the establishment and/or maintenance of normal patterns of ZPA gene expression.

The available evidence suggests that reciprocal signalling occurs between the ZPA and the AER; in particular, FGF-4 in the posterior AER and *Shh* in the ZPA are required to maintain each other's expression²⁶⁻²⁹. We detect a variable loss of *Fgf-4* expression in *Wnt-7a*^{-/-} limbs that parallels the loss of *Shh* expression in the ZPA (Fig. 4f, g). The strongest remnants of *Fgf-4* expression are found at the posterior margin of the normal *Fgf-4* domain, adjacent to the site of residual *Shh* expression. In

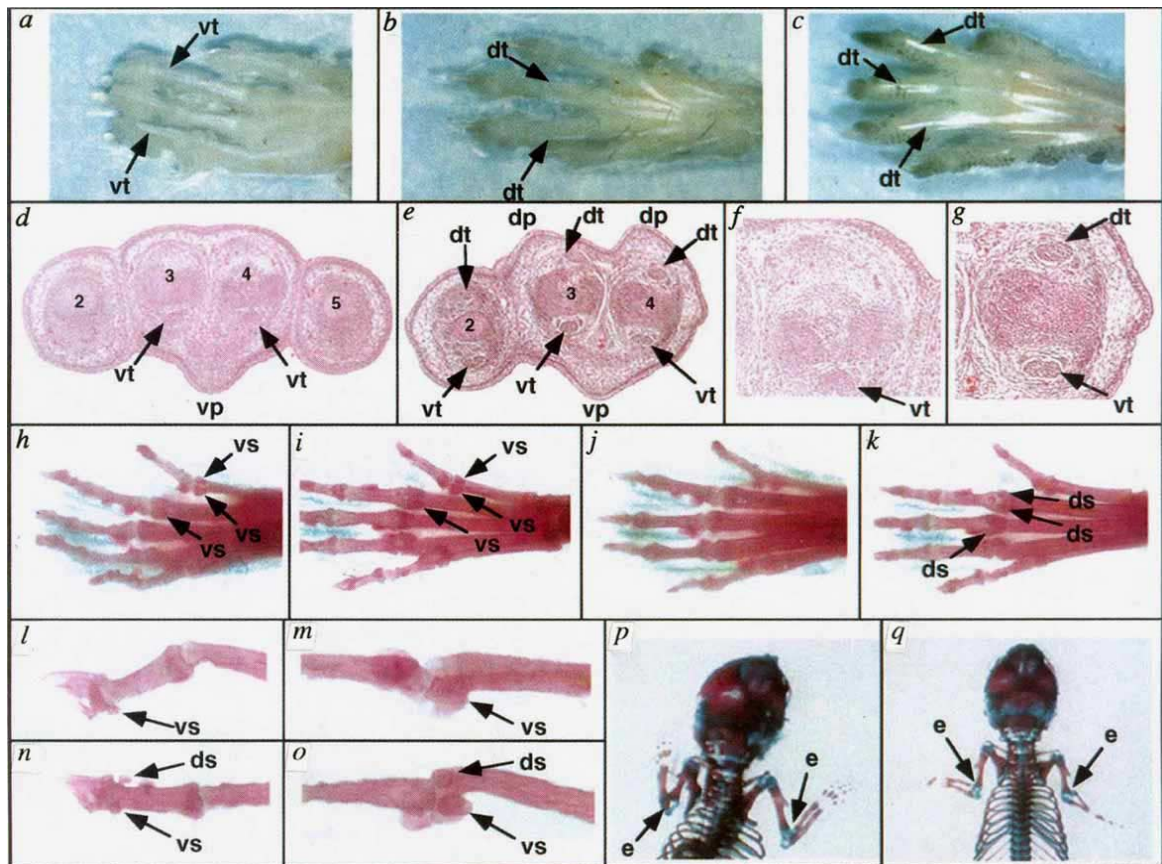


FIG. 3 Dorsal to ventral transformations of cell fate in *Wnt-7a* mutant limbs. *a-c*, Adult tendons showing wild-type ventral (*a*) and dorsal (*b*) and mutant dorsal (*c*) patterns. The dorsal tendons (*dt*) of the mutant limbs closely resemble normal ventral tendons (*vt*). *d-g*, Histological sections stained with haematoxylin and eosin through 15.5 d.p.c. forelimbs of normal (*d, f*) and mutant (*e, g*) embryos. The ventral footpads (*vp* in *d*) are duplicated on the dorsal half of mutant limbs (panel *e*). *f, g*, Higher magnification views of digit 4 from the same sections illustrate the duplication of ventral tendons (*vt*) in the dorsal half of the mutant limbs (*dt*). *h-k*, Sesamoid bones (*vs*) are present at the end of the metacarpals in the ventral half of wild type (*h*) and mutant (*i*) forelimbs. No dorsal sesamoids (*ds*) are present in wild-type forelimbs (*j*), but ectopic dorsal sesamoids are observed in mutant limbs (*k*). *l-o*, Higher-magnification view illustrating duplication of sesamoids at the

distal end of the phalanges (*n*) and metacarpals (*o*) of mutant limbs compared to wild type (*l, m*). Note the decreased flexion of mutant digits (*n*) compared to wild type (*l*). *p, q*, Abnormal flexion of mutant forelimbs of neonates (*q*) relative to wild type (*p*) at the wrist and elbow (*e*) joints.

METHODS. The skeletons of newborn pups were fixed and stained by a modification of previous procedures. Pups were skinned, eviscerated, fixed in ethanol overnight, and incubated in acetone for 24 h. They were stained in 85 ml 70% ethanol, 5 ml glacial acetic acid, 5 ml 0.3% alcian blue in 70% ethanol and 5 ml 0.1% alizarin red in 95% ethanol for 4-6 h at 37 °C, followed by 3 days at room temperature. The skeletons were cleared in 1% KOH for 1 day, then passed through a glycerol series (20, 50, 80, and 100%) of variable duration. The length of individual steps in the protocol was increased for older animals.

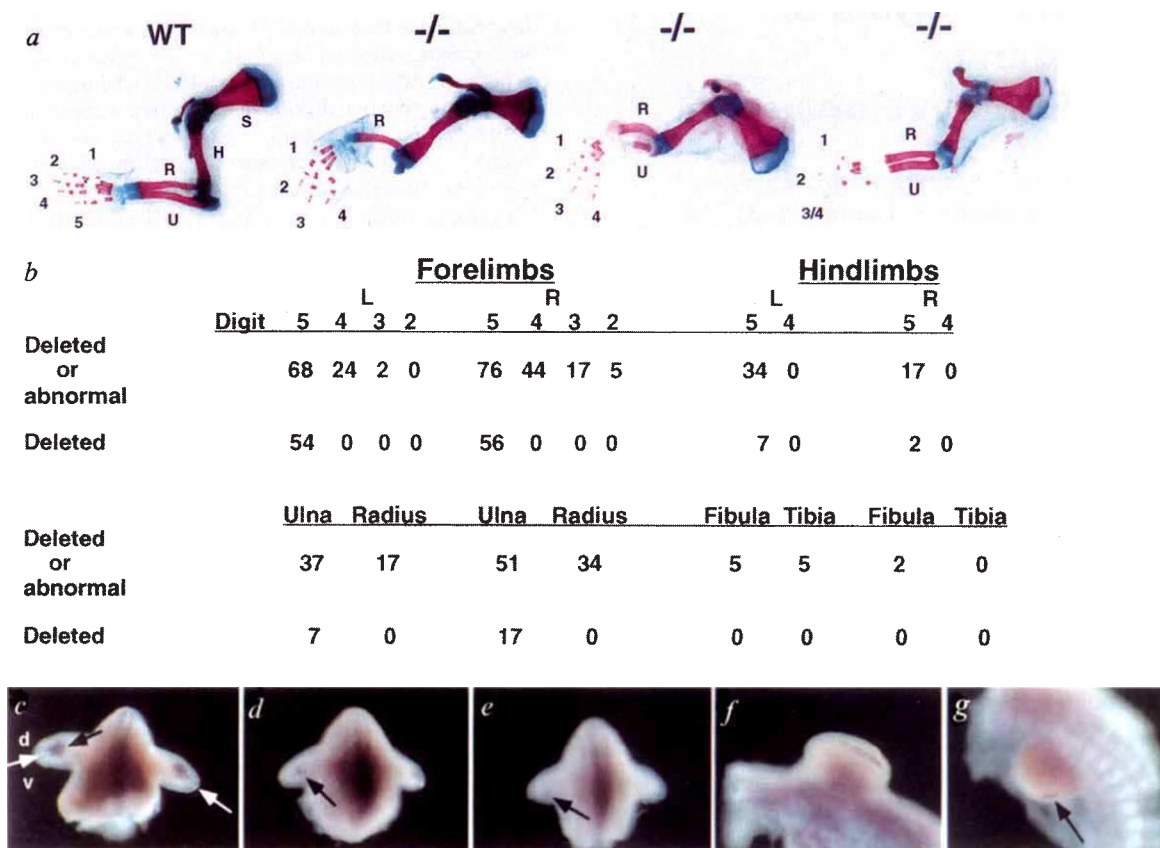


FIG. 4 *Wnt-7a* homozygotes preferentially lose posterior skeletal elements. **a**, Skeletal preparations of neonatal wild-type and mutant forelimbs demonstrate the loss or reduction of posterior digits and the ulna and a range of observed phenotypes. S, scapula; H, humerus; R, radius; U, ulna; 1–5, digits 1–5. **b**, Tabulation of the loss of skeletal structures. Numbers represent percentage of abnormal development in postnatal *Wnt-7a* mutant limbs ($n=41$). Skeletal preparations of mutant limbs were compared to wild type and scored as abnormal if they fell outside the range observed in wild-type specimens. **c–g**, Changes in *Shh* and

Fgf-4 gene expression in *Wnt-7a* mutant forelimbs at 10.25 d.p.c. **c–e**, Posterior view of *Shh* expression in wild-type (**c**) and mutant (**d**, **e**) limb buds. Wild-type *Shh* expression is stronger in the dorsal half of the posterior mesenchyme (arrows denote the dorsal–ventral midline; d, dorsal; v, ventral). *Shh* expression in mutant embryos (arrows) is substantially decreased throughout the ZPA region. **f**, **g**, *Fgf-4* expression in wild-type (**f**) and mutant (**g**) limb buds demonstrates that residual *Fgf-4* expression in mutants is concentrated at the posterior margin of the normal expression domain (arrow).

contrast, residual *Shh* expression in mutants is not localized near the *Fgf-4* domain, but is lost uniformly throughout the ZPA. Thus, *Wnt-7a* may directly regulate *Shh* expression, and the loss of *Fgf-4* may be secondary to the loss of *Shh*. In fact, *Wnt-7a* is capable of inducing *Shh* expression in the limb under experimental conditions³⁰.

Our results show that *Wnt-7a* is important in regulating both

dorsal–ventral and anterior–posterior polarity in the vertebrate limb, supporting the view that limb patterning is regulated by interactions among a network of signalling factors^{28,29}. A major challenge will be to determine the mechanisms that integrate information encoded by these various signals to produce the appropriate three-dimensional structures for particular positions in the vertebrate limb. □

Received 13 January; accepted 15 February 1995.

- Tabin, C. J. *Cell* **66**, 199–217 (1991).
- Summerbell, D., Lewis, J. H. & Wolpert, L. *Nature* **244**, 492–496 (1973).
- Saunders, J. W. *J. exp. Zool.* **108**, 363–403 (1948).
- Summerbell, D. *J. Embryol. exp. Morph.* **32**, 651–660 (1974).
- Globus, M. & Vethamany-Globus, S. *Differentiation* **6**, 91–96 (1976).
- Reiter, R. S. & Solorsh, M. *Dev. Biol.* **93**, 28–35 (1982).
- Saunders, J. W. & Gasseling, M. T. in *Epithelial–Mesenchymal Interactions* (eds Fleischmajer, R. & Billingham, R. E.) 78–97 (Williams and Wilkins, Baltimore, 1968).
- Tickle, C., Summerbell, D. & Wolpert, L. *Nature* **254**, 199–202 (1975).
- Tickle, C. *Nature* **289**, 295–298 (1981).
- Finch, R. & Zwilling, E. *J. exp. Zool.* **176**, 397–408 (1971).
- MacCabe, J. A., Errick, J. & Saunders, J. W. *Dev. Biol.* **39**, 69–82 (1974).
- Stark, R. J. & Searls, R. L. *Dev. Biol.* **38**, 51–63 (1974).
- Geduspan, J. S. & MacCabe, J. A. *Dev. Biol.* **124**, 398–408 (1987).
- Geduspan, J. S. & MacCabe, J. A. *Anat. Rec.* **224**, 79–87 (1989).
- Niswander, L., Tickle, C., Vogel, A., Booth, I. & Martin, G. R. *Cell* **75**, 579–587 (1993).
- Fallon, J. F. et al. *Science* **264**, 104–107 (1994).
- Echelard, Y. et al. *Cell* **75**, 1417–1430 (1993).
- Riddle, R. D., Johnson, R. L., Laufer, E. & Tabin, C. *Cell* **75**, 1401–1416 (1993).

- Parr, B. A., Shea, M. J., Vassileva, G. & McMahon, A. P. *Development* **119**, 247–261 (1993).
- Dealy, C. N., Roth, A., Ferrari, D., Brown, A. M. C. & Koshier, R. A. *Mech. Dev.* **43**, 175–186 (1993).
- Lyons, K. M., Pelton, R. W. & Hogan, B. L. M. *Development* **109**, 833–844 (1990).
- Davis, C. A., Holmyard, D. P., Millen, K. J. & Joyner, A. L. *Development* **111**, 287–298 (1991).
- Bulfone, A. et al. *Mech. Dev.* **40**, 129–140 (1993).
- Smith, J. C., Tickle, C. & Wolpert, L. *Nature* **272**, 612–613 (1978).
- Lee, J. J. et al. *Science* **266**, 1528–1537 (1994).
- Anderson, R., Landry, M. & Muneoka, K. *Development* **117**, 1421–1433 (1993).
- Vogel, A. & Tickle, C. *Development* **119**, 199–206 (1993).
- Niswander, L., Jeffrey, S., Martin, G. R. & Tickle, C. *Nature* **371**, 609–612 (1994).
- Laufer, E., Nelson, C. E., Johnson, R. L., Morgan, B. A. & Tabin, C. *Cell* **79**, 993–1003 (1994).
- Yang, Y. & Niswander, L. *Cell* (in the press).
- Swiatek, P. J. & Gridley, T. *Genes Dev.* **7**, 2071–2084 (1993).

ACKNOWLEDGEMENTS. We thank J. McMahon for injecting blastocysts; B. Klumppar for sectioning; T. Gridley for C7 ES cells; Y. Yang and L. Niswander for communicating results before publication; R. Balling, A. Joyner, K. Lyons, L. Niswander and J. Rubenstein for probes and antibodies; and J. Saunders, L. Niswander, C. Tickle, D. Summerbell, J. Smith and our colleagues for discussion. This work was supported by a grant from the NIH to A.P.M.

Functional analysis of activins during mammalian development

Martin M. Matzuk^{*†‡}, T. Rajendra Kumar[†],
Anne Vassalli[§], Jackie R. Bickenbach[‡],
Dennis R. Roop[‡], Rudolf Jaenisch[§]
& Allan Bradley^{*¶}

Departments of ^{*} Molecular and Human Genetics, [†] Pathology and [‡] Cell Biology, [¶] Howard Hughes Medical Institute, Baylor College of Medicine, Houston, Texas 77030, USA
[§] Whitehead Institute for Biomedical Research, Cambridge, Massachusetts 02142, USA ^{||} Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

ACTIVINS are dimeric ($\beta\alpha\beta\alpha$; $\beta\beta\beta\beta$; $\beta\alpha\beta\beta$) members of the transforming growth factor- β superfamily¹. They are widely expressed during murine development¹⁻⁶, are highly conserved during vertebrate evolution^{1,7-11}, and may be involved in mesoderm induction and neurulation in *Xenopus laevis* and *Oryzias latipes*¹⁰⁻¹⁷. To investigate the function of activins during mammalian development, we generated mice with mutations either in activin- $\beta\alpha$ or in both activin- $\beta\alpha$ and activin- $\beta\beta$. Activin- $\beta\alpha$ -deficient mice develop to term but die within 24 h of birth. They lack whiskers and lower incisors

and have defects in their secondary palates, including cleft palate, demonstrating that activin- $\beta\alpha$ must have a role during craniofacial development. Mice lacking both activin subunits show the defects of both individual mutants but no additional defects, indicating that there is no functional redundancy between these proteins during embryogenesis. In contrast to observations in lower vertebrates¹⁰⁻¹⁷, zygotic expression of activins is not essential for mesoderm formation in mice.

In the mouse, activin- $\beta\alpha$ and $\beta\beta$ subunits (Fig. 1a) are expressed zygotically before implantation^{2,4}. After implantation, activin subunit expression is limited to the maternal deciduum until E10.5, when activin- $\beta\alpha$ begins to be expressed in mesenchymal cells of the developing face, whiskers, hair follicles, heart and digestive tract^{5,6,18}. Mice deficient in activin- $\beta\beta$ are viable but have defective eyelid development and female reproduction and normal mesoderm formation or neurulation¹⁹. To investigate the function of activins during mammalian development, mice with a disrupted activin- $\beta\alpha$ allele were generated using embryonic stem cell technology (Fig. 1b). Heterozygous (*act $\beta\alpha$ ^{ml}/+*) mice were fertile and viable and were intercrossed to obtain *act $\beta\alpha$ ^{ml}/act $\beta\alpha$ ^{ml}* mice (activin- $\beta\alpha$ -deficient). Genotype analysis of the progeny from these intercrosses (Fig. 1c) revealed normal mendelian ratios, with 25.1% wild type (68/271), 49.4% heterozygotes (134/271) and 25.5% homozygotes (69/271), indicating that *act $\beta\alpha$ ^{ml}/act $\beta\alpha$ ^{ml}* mice survived to birth. The activin- $\beta\alpha$ -deficient mice (Fig. 2a) appeared to be healthy at birth, breathed normally, and were similar in weight to controls (*act $\beta\alpha$ ^{ml}/act $\beta\alpha$ ^{ml}*, 1.19 ± 0.12 g ($n = 22$); littermate controls, 1.24 ± 0.11 g ($n = 23$)). Although (*act $\beta\alpha$ ^{ml}/act $\beta\alpha$ ^{ml}*

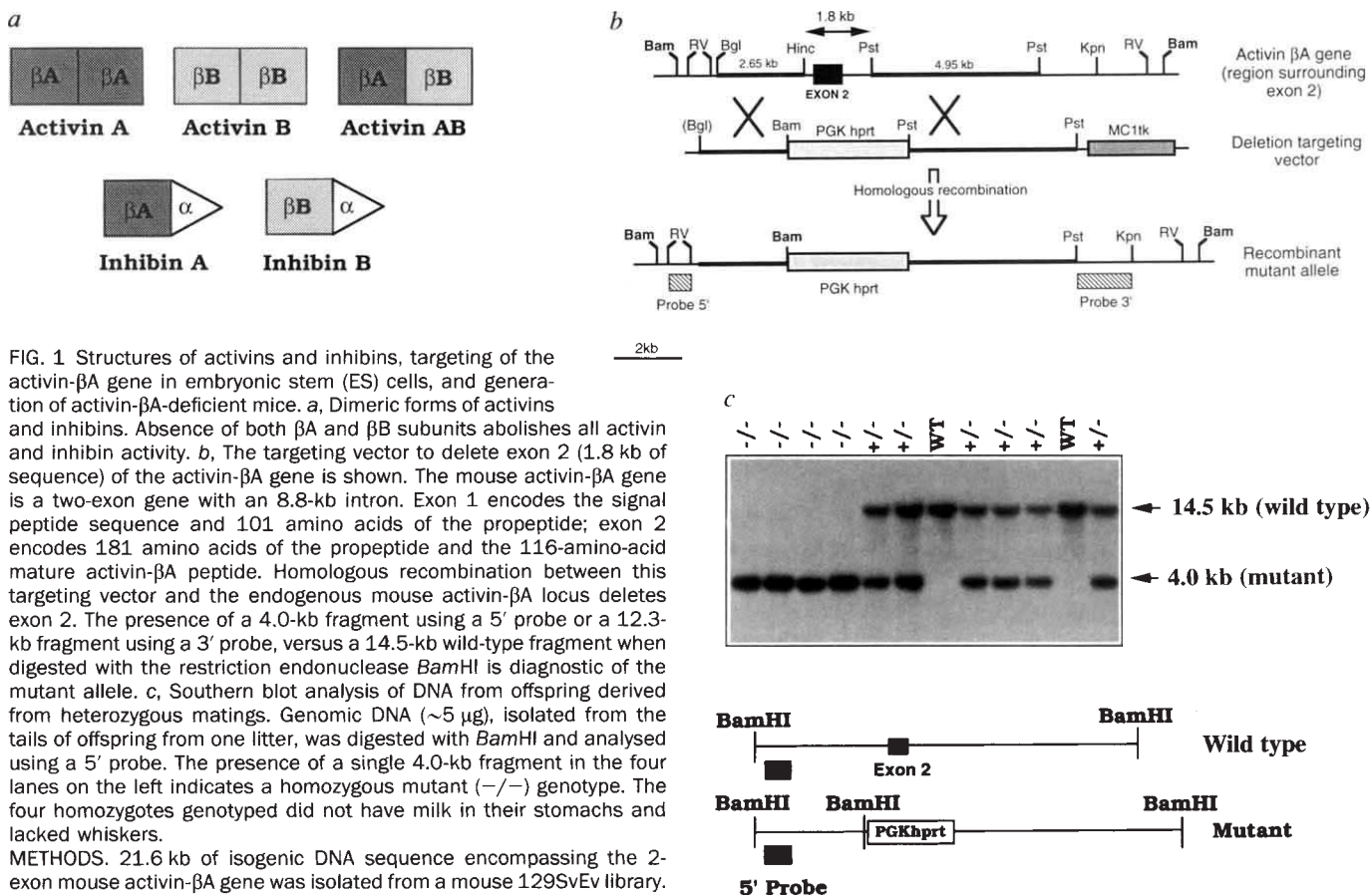


FIG. 1 Structures of activins and inhibins, targeting of the activin- $\beta\alpha$ gene in embryonic stem (ES) cells, and generation of activin- $\beta\alpha$ -deficient mice. **a**, Dimeric forms of activins and inhibins. Absence of both $\beta\alpha$ and $\beta\beta$ subunits abolishes all activin and inhibin activity. **b**, The targeting vector to delete exon 2 (1.8 kb of sequence) of the activin- $\beta\alpha$ gene is shown. The mouse activin- $\beta\alpha$ gene is a two-exon gene with an 8.8-kb intron. Exon 1 encodes the signal peptide sequence and 101 amino acids of the propeptide; exon 2 encodes 181 amino acids of the propeptide and the 116-amino-acid mature activin- $\beta\alpha$ peptide. Homologous recombination between this targeting vector and the endogenous mouse activin- $\beta\alpha$ locus deletes exon 2. The presence of a 4.0-kb fragment using a 5' probe or a 12.3-kb fragment using a 3' probe, versus a 14.5-kb wild-type fragment when digested with the restriction endonuclease *Bam*HI is diagnostic of the mutant allele. **c**, Southern blot analysis of DNA from offspring derived from heterozygous matings. Genomic DNA (~5 μ g), isolated from the tails of offspring from one litter, was digested with *Bam*HI and analysed using a 5' probe. The presence of a single 4.0-kb fragment in the four lanes on the left indicates a homozygous mutant (-/-) genotype. The four homozygotes genotyped did not have milk in their stomachs and lacked whiskers.

METHODS. 21.6 kb of isogenic DNA sequence encompassing the 2-exon mouse activin- $\beta\alpha$ gene was isolated from a mouse 129SvEv library. Linearized vector (25 μ g) was electroporated into the *hprt*-negative AB2.1 ES cell line, selected in HAT (where HAT is hypoxanthine, aminopterin, thymidine) and FIAU is (1-(2'-deoxy-2'-fluoro- β -D-arabinofuranosyl)-5-iodouracil), analysed by Southern blot analysis, and injected into chimaeras as described^{30,31}. Enrichment in HAT and FIAU was 14.4-fold compared to HAT alone. Southern blot analysis of ES cell DNA identified 14 targeted clones out of 115 clones screened. The mutant

allele was transmitted from chimaeras derived from independent clones $\beta\alpha$ 5-F12 and $\beta\alpha$ 6-D5. To confirm the nature of the mutant allele, a cDNA probe encoding the activin- $\beta\alpha$ C-terminal (mature) sequence was hybridized to DNA from the *act $\beta\alpha$ ^{ml}/act $\beta\alpha$ ^{ml}* mice. The lack of hybridization to a wild-type activin- $\beta\alpha$ allele confirmed that the mutant *act $\beta\alpha$ ^{ml}* allele was null (data not shown).

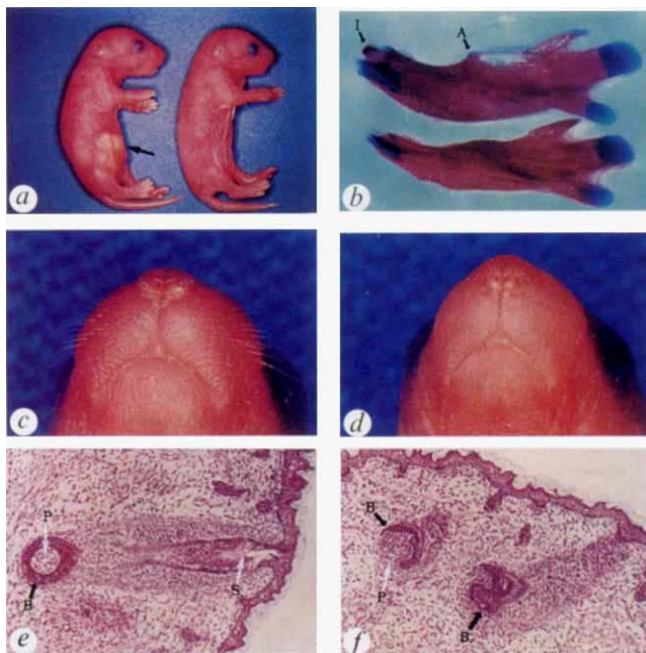


FIG. 2 Morphological and histological analysis of activin- β A mutant and control mice. **a**, Wild-type (left) and activin- β A-deficient homozygote (right) newborn littermate mice. Note the absence of milk in the abdomen of the activin- β A-deficient mouse compared to the wild-type control (arrow). **b**, Medial view of wild-type (top) and $act\beta A^{ml}/act\beta A^{ml}$ (bottom) dissected mandibles stained with alizarin red and alcian blue. The alveolar ridge (A) is the region of the mandible that surrounds the lower molar teeth. The alveolar ridge (A) is less prominent and the incisor (I) is missing in the mutant, consistent with the expression of activin β A in the developing tooth buds^{6,18}. There are no primary defects in mandible development. **c** and **d**, Gross analysis of the whisker pads and face of a wild-type control (**c**) and $act\beta A^{ml}/act\beta A^{ml}$ (**d**) newborn mice, showing the absence of whiskers in the mutant (**d**) mouse. **e** and **f**, Histological analysis (haematoxylin and eosin stain³¹) of the whisker follicles of control (**e**) and $act\beta A^{ml}/act\beta A^{ml}$ (**f**) mice; magnification is the same in **e** and **f** and sections are transverse sections through the whisker pads. Note that the distances of the hair papillae (P) from the surface in the mutant (**f**) are shortened compared to the control (**e**). In addition, the hair bulbs (B) in the mutants (**f**) are less well formed than in the control (**e**). The root sheath (S) is obvious as it approaches the epidermis in the control (**e**), but is rarely visible in sections taken from several mutants.

mice progressed to term at the expected frequency, they died within 24 h of birth. The $act\beta A^{ml}/act\beta A^{ml}$ mice lacked whiskers and incisors (Fig. 2b–d), consistent with the expression of activin- β A messenger RNA in the mesenchyme of the whisker follicle and teeth primordia^{6,18}. Analysis of skeletal preparations showed that the lower incisors were absent (complete absence in 22/25 $act\beta A^{ml}/act\beta A^{ml}$ mice) and that there was a secondary defect in the alveolar ridge of the mandible, the site of formation of the lower molars (Fig. 2b). Immature tooth buds could be detected in histological sections (Fig. 3). Histological analysis of the whisker pads revealed that growth and/or differentiation of the whisker follicles was delayed (Fig. 2e, f). The root sheath, if present, appears to be immature, the cells at the base of the hair bulb often form an irregular pattern, and the hair papilla always lies closer to the surface (Fig. 2f); none of these features is evident in control littermate mice. Thus, activin A is required for the normal development of whiskers and teeth.

Newborn $act\beta A^{ml}/act\beta A^{ml}$ mice, identified by their lack of whiskers, failed to suckle because of a cleft palate; 12 of 42 (29%) hybrid background (CS7B1/6/129SvEv) and 3 of 9 (33%) 129SvEv inbred $act\beta A^{ml}/act\beta A^{ml}$ newborn mice had a cleft secondary palate (Fig. 3). Skeletal preparations of 22 $act\beta A^{ml}/$

$act\beta A^{ml}$ newborn mice (hybrid background) that did not have clefts of the secondary palate revealed that 7 of 22 (32%) lacked a hard palate, whereas in most of the remaining mice development of the hard palate was incomplete (Fig. 3).

To generate mice deficient in all activins and therefore lacking both activin- β A and β B subunits (Fig. 1a), $act\beta A^{ml}/+$ mice were interbred with $act\beta B^{ml}/+$ mice¹⁹ to generate compound heterozygous mutant mice ($act\beta A^{ml}, act\beta B^{ml}/+, +$). Compound heterozygous mutant mice were interbred and compound homozygous mutant mice ($act\beta A^{ml}, act\beta B^{ml}/act\beta A^{ml}, act\beta B^{ml}$) deficient in all activins and inhibins were viable at birth (Fig. 3f) and seen at the expected frequency (Table 1). $act\beta A^{ml}, act\beta B^{ml}/act\beta A^{ml}, act\beta B^{ml}$ mice were readily identifiable as they exhibited phenotypic characteristics of the individual mutant lines. These mice lacked whiskers and incisors (data not shown) and had eyelid defects (namely open eyes; Fig. 3f), and appear to have died of palate defects, with three out of nine (33%) having cleft secondary palate. Furthermore, the weight of $act\beta A^{ml}, act\beta B^{ml}/act\beta A^{ml}, act\beta B^{ml}$ mice was comparable (1.08 ± 0.11 g; $n=9$) to their littermates (1.08 ± 0.13 g; $n=35$) and they had no additional defects over those seen in individual activin- β A or β B mutant mice.

The absence of whiskers and incisors and the defects in the formation of the secondary palate in activin- β A-deficient mice

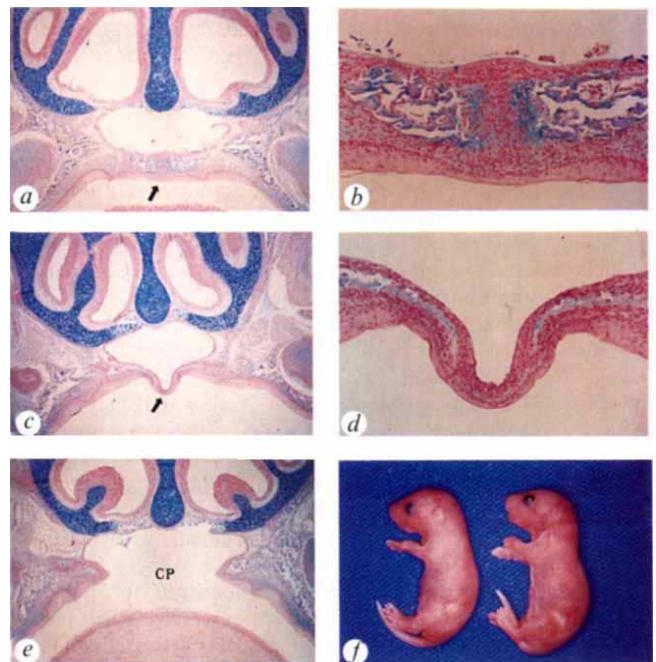


FIG. 3 Histological analysis of the palate and surrounding tissues (**a–e**) and morphological analysis of compound homozygote mice (**f**). Coronal sections at the level of the eye of wild-type (**a, b**) and $act\beta A^{ml}/act\beta A^{ml}$ (**c–e**) mice. Note the normal joining of the lateral palatine processes to form the hard and soft palates in the wild-type mouse (arrow) photographed at low (**a**) and high (**b**) magnification. The hard palate (arrow) has not formed in the mutant at low (**c**) and high (**d**) magnification, which is evident from a lack of bone matrix and a sagging of the palate. There is a cleft palate (CP; absence of both hard and soft palates) in another mutant (**e**). **f**, Newborn control (left) and compound homozygote $act\beta A^{ml}, act\beta B^{ml}/act\beta A^{ml}, act\beta B^{ml}$ (right) mice. Both mice were alive at the time the photograph was taken. The mutant mouse (right) had open eyes and no whiskers and was approximately the same size and weight as the control.

METHODS. Coronal sections through the palates were processed as described³¹ and stained with alcian blue and neutral red. Southern blots of DNA from newborn offspring generated by intercrossing of compound heterozygous activin-mutant mice were analysed as described^{19,31}, using the 5' probe to determine the activin- β A genotype and a 3' probe for the activin- β B genotype¹⁹. Tail DNA was digested with *Bam*HI.

TABLE 1 Genotype analysis of offspring generated from compound heterozygote mice

βA	βB	Number	Observed (%)	Expected (%)
WT*	WT	10	8.0	6.25
WT	+/-	13	10.4	12.5
WT	-/-	10	8.0	6.25
+/-	WT	11	8.8	12.5
+/-	+/-	34	27.2	25
+/-	-/-	16	12.8	12.5
-/-	WT	4	3.2	6.25
-/-	+/-	13	10.4	12.5
-/-	-/-	14	11.2	6.25
Total		125	100.0	100.0

* WT, wild type; +/-, heterozygote; -/-, homozygote.

are consistent with the expression of activin- βA mRNA^{6,18}, and with the *in vitro* effects of activins on chondrogenesis and bone formation^{20,21}. The cleft palate and mandible and incisor defects are also seen in *Mxsl*-deficient mice²². The incomplete penetrance of the 'cleft' is reminiscent of human cleft palate and the multifactorial nature of its development²³. Interestingly, mice with the recessive mutation, *oel*, had cleft palate, open eyelids, and died perinatally²⁴, like compound homozygous activin- $\beta A/\beta B$ mutant mice. The phenotype of *oel* mice is consistent with a defect in a common downstream component of both the activin A and B signal transduction cascades. The observation that activin- $\beta A/\beta B$ double-mutant mice have a phenotype that is an additive combination of the single mutant phenotypes indicates that individual activin homodimers are not functionally compensating for one another during development and that the activin- $\beta A\beta B$ heterodimer does not have a unique function during embryogenesis.

Although we have shown that zygotic activins are not essential for mesoderm induction in mice, experiments with mice carrying a TGF- $\beta 1$ mutation indicate that maternal TGF- $\beta 1$ may cross the placenta to partially rescue the mutant mice²⁵. As activin- βB -deficient mice have normal mesoderm-derived tissue when delivered from activin- βB -deficient mothers¹⁹, maternal or decidua activin B is not essential; similar studies with activin- βA -deficient females are not possible. Overexpression of activin subunits, receptors and truncated receptors in *Xenopus laevis* and zebra fish could influence mesoderm induction¹⁰⁻¹⁷ as a result of interaction with other TGF- β superfamily members and their receptors²⁶⁻²⁸. For example, truncated *Xenopus* activin type-II receptors (ActRcII or ActRcIIB) block the activity of Vg1, a related TGF- β superfamily member²⁸. Identification of the serine/threonine kinase receptors that are expressed at or around the time of mesoderm induction will give us a better understanding of the TGF- β -related proteins involved in this process. Lastly, the phenotype of mutant mice deficient in one of the activin receptors, ActRcII (see accompanying Letter²⁹), overlaps minimally with that of activin-deficient mice. This calls into question the relevance of ActRcII as a receptor for activin A and/or activin B during embryogenesis. Alternatively, the other type-II activin receptor, ActRcIIB, might be important in activin signal transduction during embryogenesis. A full understanding of the specificity of these ligand/receptor interactions must come from analysis of ActRcIIB mutant mice and the demonstration of epistasis in compound mutant ligand/receptor mice. □

Received 9 November 1994, accepted 24 January 1995.

- Vale, W., Hsueh, A., Rivier, C. & Yu, J. *Peptide Growth Factors and their Receptors II* (eds Sporn, M. B. & Roberts, A. B.) (Springer, Berlin, 1990).
- Manova, K., Paynton, B. V. & Bachvarova, R. F. *Mech. Dev.* **36**, 141-152 (1992).
- Albano, R. M., Groome, N. & Smith, J. C. *Development* **117**, 711-723 (1993).
- Lu, R.-Z., Shigemitsu, M., Nishihara, M. & Takahashi, M. *Biol. Reprod.* **49**, 1163-1169 (1993).
- Albano, R. M., Arkell, R., Beddington, R. S. P. & Smith, J. C. *Development* **120**, 803-813 (1994).
- Feijen, A., Goumans, M. J. & van den Eijnden-van Raaij, A. J. M. *Development* **120**, 3621-3637 (1994).

- Esch, F. S. et al. *Molec. Endocr.* **1**, 388-396 (1987).
- Mitrani, E. et al. *Cell* **63**, 495-501 (1990).
- Ge, W., Gallin, W. J., Strobeck, C. & Peter, R. E. *Biochem. biophys. Res. Commun.* **193**, 711-717 (1993).
- Smith, J. C., Price, B. M. J., Van Nimmen, K. & Huylebroeck, D. *Nature* **345**, 729-731 (1990).
- Thomsen, G. et al. *Cell* **63**, 485-493 (1990).
- Kondo, M. *Biochem. biophys. Res. Commun.* **181**, 684-690 (1991).
- Hemmati-Brivanlou, A. & Melton, D. A. *Nature* **359**, 609-614 (1992).
- Mathews, L. S., Vale, W. E. & Kintner, C. R. *Science* **255**, 1702-1705 (1992).
- Hemmati-Brivanlou, A. & Melton, D. A. *Cell* **77**, 273-281 (1994).
- Wittbrodt, J. & Rosa, F. M. *Gene Dev.* **8**, 1448-1462 (1994).
- Gurdon, J. B., Harger, P., Mitchell, A. & Lemaire, P. *Nature* **371**, 487-492 (1994).
- Roberts, V. J. & Barth, S. *Endocrinology* **134**, 914-923 (1994).
- Vassalli, A., Matzuk, M. M., Gardner, H. A. R., Lee, K.-F. & Jaenisch, R. *Genes Dev.* **8**, 414-427 (1994).
- Centrella, M., McCarthy, T. L. & Canalis, E. *Molec. cell. Biol.* **11**, 250-258 (1991).
- Luyten, F. P., Chen, P., Paralkar, V. & Reddi, A. H. *Exp. Cell Res.* **210**, 224-229 (1994).
- Satokata, I. & Maas, R. *Nature Genet.* **6**, 348-356 (1994).
- Gorlin, R. J., Cohen, M. M. & Levin, L. S. *Syndromes of the Head and Neck* (Oxford Monogr. Med. Genet. No. 19) 3rd ed 693-714 (Oxford University Press, Oxford, UK, 1990).
- Gluecksohn-Waelsch, S. *Mouse Newslett.* **25**, 12 (1961).
- Letterio, J. J. et al. *Science* **264**, 1936-1938 (1994).
- Zhou, X., Sasaki, H., Lowe, L., Hogan, B. L. M. & Kuehn, M. R. *Nature* **361**, 543-547 (1993).
- Conlon, F. et al. *Development* **120**, 1919-1928 (1994).
- Schulte-Merker, S., Smith, J. C. & Dale, L. *EMBO J.* **13**, 3533-3541 (1994).
- Matzuk, M. M., Kumar, T. R. & Bradley, A. *Nature* **374**, 356-360 (1995).
- Bradley, A. in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach* (ed. Robinson, E. J.) (IRL, Oxford, 1987).
- Matzuk, M. M., Finegold, M. J., Su, J.-G.J., Hsueh, A. J. W. & Bradley, A. *Nature* **360**, 313-319 (1992).

ACKNOWLEDGEMENTS. We thank N. Lu for technical assistance; W. Moyle and R. Myers for help in isolating activin- βA cDNA; M. Nulley for isolating bacteriophage harbouring the mouse activin- βA gene; A. Roe for help in constructing activin- βA targeting vector; S. Baker and B. Powell for manuscript preparation; and R. Behringer for critically reviewing the manuscript. This research was supported in part by NIH grants (to M.M.M.). A.B. is an Associate Investigator with the Howard Hughes Medical Institute.

Different phenotypes for mice deficient in either activins or activin receptor type II

Martin M. Matzuk*†‡, T. Rajendra Kumar† & Allan Bradley*§

Departments of * Molecular and Human Genetics, † Pathology and ‡ Cell Biology, § Howard Hughes Medical Institute, Baylor College of Medicine, Houston, Texas 77030, USA

ACTIVINS are believed to initiate a signal transduction cascade by binding to serine/threonine kinase receptors types I and II¹⁻³. Activins⁴ bind to several different receptors *in vitro*¹⁻³, but the significance of this interaction *in vivo* has not been confirmed. To test the function of the type II activin receptor (ActRcII) in mammalian development and reproduction, we generated a null mutation in the ActRcII gene in mice using embryonic stem cell technology. We expected ActRcII-deficient mice to phenocopy activin-deficient mice. A few ActRcII-deficient mice had skeletal and facial abnormalities reminiscent of the Pierre-Robin syndrome in humans^{5,6}, but most lacked these defects and developed into adults; their follicle-stimulating hormone was suppressed, and their reproductive performance was defective. These findings confirm a role of ActRcII in activin signalling in pituitary gonadotrophs. The striking lack of overlap between phenotypes of ActRcII-deficient and activin-deficient mice suggests that the ligands that signal through ActRcII during embryonic development are not activins.

We tested the function of ActRcII in activin signalling by mutating the ActRcII gene using ES cell technology to delete exon 1, which includes both the initiation codon and the signal peptide sequence (Fig. 1a). Mice heterozygous for this null allele (*actRcII*^{ml}/+) were intercrossed to produce homozygote F₂ ActRcII-deficient (*actRcII*^{ml}/*actRcII*^{ml}) mice (Fig. 1b, c). ActRcII-deficient mice would be expected to be a phenocopy of the activin- βA , activin- βB or activin- $\beta A/\beta B$ double-mutant mice (that is, to exhibit defects in eyelid development, lack whiskers

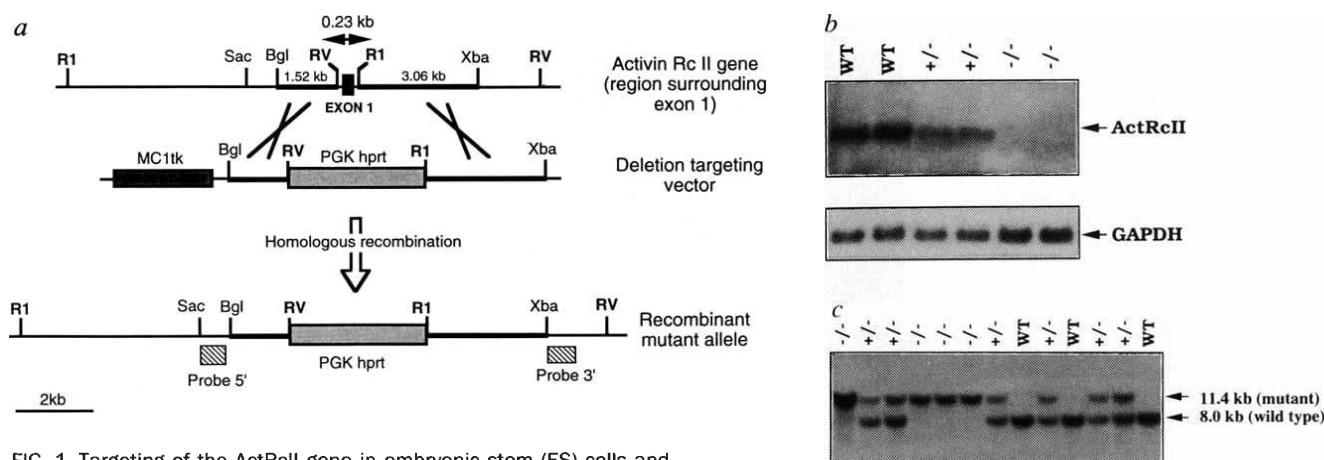


FIG. 1 Targeting of the *ActRcII* gene in embryonic stem (ES) cells and Southern and northern blot analysis of DNA and RNA from offspring derived from heterozygote intercrosses. **a**, The targeting vector to delete exon 1 (0.23 kb of sequence) of the *ActRcII* gene is shown. The *ActRcII* gene is large (>60 kb)²² and we therefore constructed a replacement vector to delete 0.23 kb of *ActRcII* gene sequence. This mutation deletes the coding sequence of exon 1 encompassing both the initiation ATG codon and the signal peptide sequence²². As the signal peptide is required for the protein to translocate to the cell surface, absence of this sequence will result in *ActRcII* deficiency. The targeting vector was electroporated into the *hprt*-negative AB2.1 line of ES cells and 8% of the ES cell clones (30 out of 369) screened by Southern blotting were correctly targeted. Germline transmission from chimaeric males generated with 5 of 6 ES cell clones was achieved. The targeting vector contains 1.52 kb of sequence upstream of exon 1, 3.06 kb of sequence 3' of exon 1, the *PGK-hprt* (where *PGK* is the phosphoglycerate kinase I promoter) expression cassette, and an *MC1-tk* (thymidine kinase) expression cassette. Homologous recombination between this targeting vector and the endogenous *ActRcII* locus (top) would enlarge the locus by 3.4 kb. The recombinant allele can be detected by restriction endonuclease digestion of ES cell DNA with either *EcoRI* (RI) or *EcoRV* (RV) and Southern blot analysis using 5' or 3' external probes, respectively. The presence of an 11.4-kb mutant allele versus a wild-type 8.0-kb fragment is diagnostic of the crossover event on the 5' arm (panel b). The presence of an 8.3-kb fragment versus a 4.9-kb wild-type fragment upon restriction with *EcoRV* is diagnostic of the mutant allele when using a 3' external probe. **b**, Genomic DNA (~5 µg), isolated from the tails of offspring from two litters, was digested with *EcoRI* and analysed

as described²³ using a 5' probe. The presence of a single 11.4-kb fragment indicates a homozygote (—/—) genotype. Genotype analysis of 555 F₂ hybrid progeny from *actRcII*^{ml}/+ × *actRcII*^{ml}/+ matings showed the following: +/+, 166 (29.5%); *actRcII*^{ml}/+, 292 (52.6%); and *actRcII*^{ml}, 97 (17.5%). Analysis of 163 F₂ 129Sv-inbred progeny from an equivalent cross demonstrated the following: +/+, 46 (28.2%); *actRcII*^{ml}/+, 86 (52.8%); *actRcII*^{ml}/*actRcII*^{ml}, 31 (19.0%). **c**, 20 µg total RNA, isolated from the livers of 6 independent age-matched adult male mice, was electrophoresed on a formaldehyde gel, transferred to Hybond-N nylon membranes (Amersham), and hybridized with either a 3' untranslated mouse *ActRcII* probe²² or a glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) probe. The *ActRcII* probe detects a 6-kb mRNA species²⁴, and the *GAPDH* probe detects a 1.4-kb mRNA species. The 6-kb *ActRcII* transcript could not be detected in livers from *actRcII*^{ml}/*actRcII*^{ml} (—/—) mice, confirming that the *ActRcII* mutation was null. Similar findings were noted by analysis of RNA derived from adult testes (data not shown). WT, wild type; +/—, heterozygote.

METHODS. The mouse *ActRcII* gene sequences were isolated from a 129SvEv genomic library as described²². Linearized vector (25 µg) was electroporated into the *hprt*-negative AB2.1 ES cell line, selected in HAT and FIAU (where HAT represents hypoxanthine, aminopterin, thymidine and FIAU is (1-(2'-deoxy-2'-fluoro-β-D-arabinofuranosyl)-5-iodouracil)) and targeted clones were identified and injected into blastocysts to generate chimaeras as described^{22,25,26}. Enrichment in HAT and FIAU was 2.7-fold compared to HAT alone.

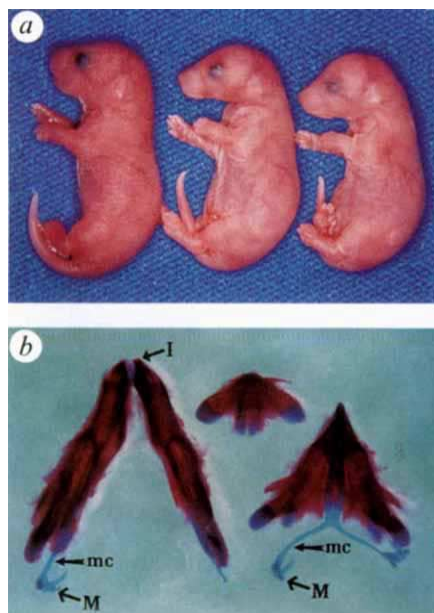
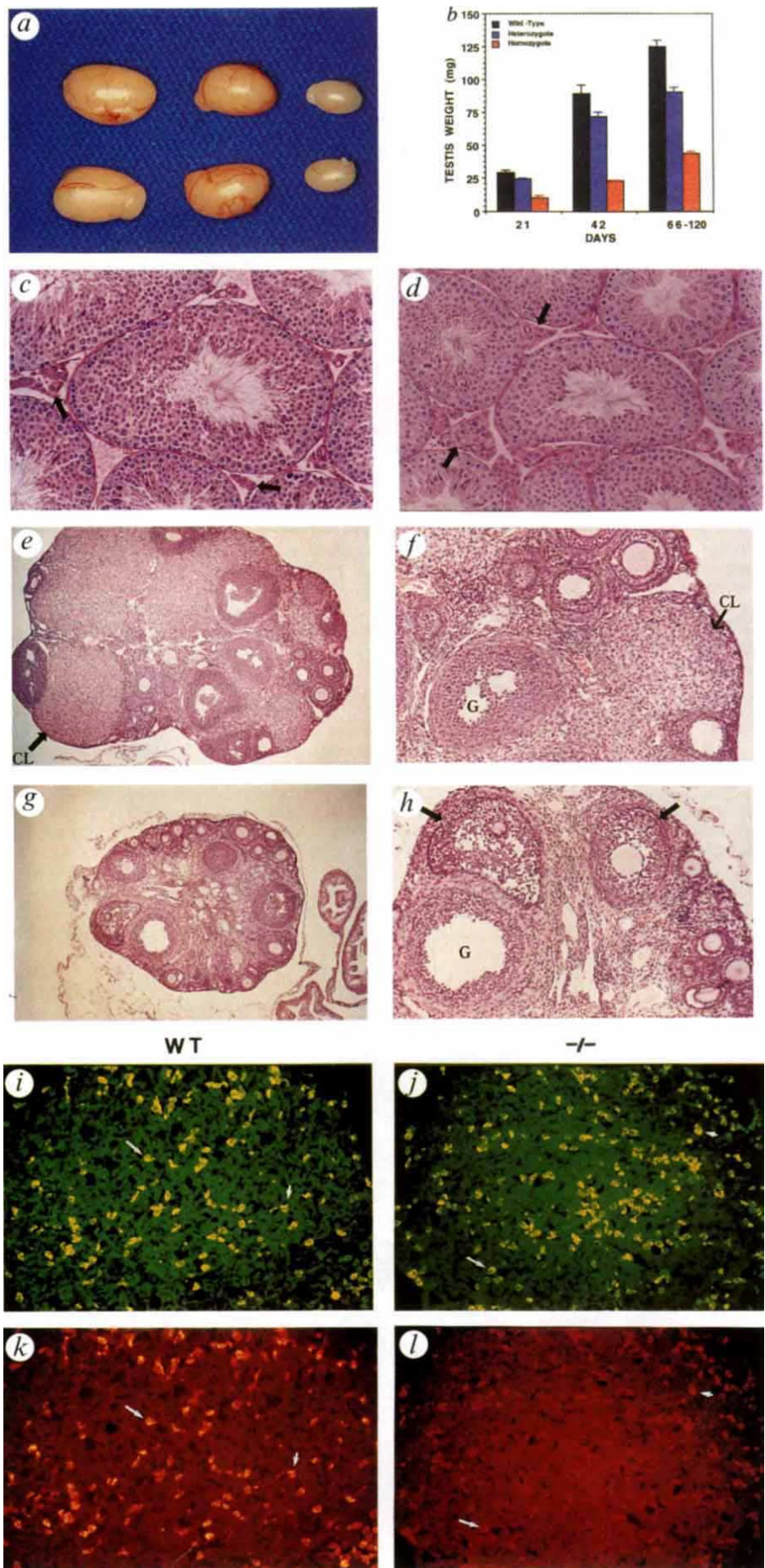


FIG. 2 Morphological analysis of E18.5 mice. **a**, Three E18.5 mice delivered from the same litter by caesarian section are shown. The mouse on the left was genotyped as *actRcII*^{ml}/+, whereas the two on the right were genotyped as *actRcII*^{ml}/*actRcII*^{ml} and died within minutes of delivery. Note that the jaws of the *actRcII*^{ml}/*actRcII*^{ml} mice are less prominent compared to the control and there are subtle eyelid closure defects. **b**, Alizarin red/alcian blue analysis²⁷ of the mandibles from these three mice are shown in the same order. The mandibles of the two *actRcII*^{ml}/*actRcII*^{ml} mice (right) are hypoplastic and are the likely cause of death for these mice. The malleus (m) is normal in the mutant on the right. Meckel's cartilage (mc) is visible but abnormal in the mutant on the right, compared to the control. The Meckel's cartilage is abnormal and not attached to the mandible in the *actRcII*^{ml}/*actRcII*^{ml} mouse in the centre. Incisors (I) are present in the normal mandible but absent in the two mandibles on the right. The mandible defects were seen in offspring derived from 2 independent ES cell lines. Skeletal defects were also seen as early as E13.5 in alcian-blue-stained embryos (data not shown).

FIG. 3 Morphology and histology of the gonads and immunofluorescent analysis of the pituitaries of *ActRcll*-deficient and control mice. **a**, Gross analysis of the testes of 42-day-old wild-type (WT; left), heterozygote (+/-; centre) and homozygote (-/-; right) littermate mice. **b**, Testis weights at different ages from 21-, 42- and 66-120-day mice. Values shown are mean $\pm$ s.e.m. The numbers (*n*) of mice used were: 21 days, WT (6), +/- (9), -/- (4); 42 days, WT (5), +/- (6), -/- (6); 66-120 days, WT (12), +/- (19), -/- (20). Testis weights from 66-120 days in each group did not vary significantly and were therefore grouped. Interestingly, *actRcll*^{ml}/+ male mice also had lighter testes, in line with the 50% decrease in *ActRcll* mRNA in these mice (Fig. 1c, and data not shown). **c** and **d**, Histology of testes from 10-week-old wild-type (**c**) and *actRcll*^{ml}/*actRcll*^{ml} (**d**) mice (same magnification). Note the decrease in tubule size in the testis from the *actRcll*^{ml}/*actRcll*^{ml} mouse (**d**) compared to the control (**c**). Stages of spermatogenesis are normal and spermatozoa are similar in the tubules of both mice. Leydig (interstitial) cells (arrows) appear to be more abundant in the testes of *actRcll*^{ml}/*actRcll*^{ml} mice (**d**), but the net number of Leydig cells in the small testes is likely to be similar to controls as pituitary luteinizing hormone (LH) levels are normal (see below), and serum testosterone levels in *actRcll*^{ml}/*actRcll*^{ml} adult mice are comparable to controls (data not shown). **e-h**, Histology of ovaries from wild-type (**e, f**) and *actRcll*^{ml}/*actRcll*^{ml} (**g, h**) mice; **e** and **g** were photographed at the same magnifications, as were **f** and **h**. Note the absence of corpora lutea (CL) in the mutant (**g, h**) versus the wild-type control (**e, f**) and the atretic follicles in the mutant ovary (arrow, **h**). Confirmation of the normal function of oocytes in *actRcll*^{ml}/*actRcll*^{ml} females is demonstrated by the response of these mice to superovulation; the ovulated eggs can be fertilized *in vivo* and develop into embryos (data not shown). **i-l**, Double-label immunofluorescence analysis of LH (**i, j**) or FSH (**k, l**) from wild-type (**i, k**) and *ActRcll*-deficient (**j, l**) male mice. The majority of the pituitary gonadotrophs in the wild-type mice (**i, k**) express both LH (**i**) and FSH (**k**). In contrast, whereas the LH content (**j**) of the *ActRcll*-deficient homozygote (-/-) mouse is comparable to the WT control (**i**), FSH (**l**) is dramatically suppressed and just above background. Arrows show the few gonadotrophs expressing both LH and FSH. Qualitative analysis of pituitary prolactin, GH, ACTH and TSH content by immunofluorescence did not reveal any differences in *actRcll*^{ml}/*actRcll*^{ml} mutants compared to controls (data not shown), as for LH. **METHODS.** Testes and ovaries were fixed in Bouin's and 10% formalin, respectively, and then processed as described²³. Testis sections were stained with haematoxylin and PAS (periodic acid Schiff); ovary sections were stained with haematoxylin and eosin. Pituitary and radioimmunoassay analyses were done as described^{23,28} (reagents from the National Hormone and Pituitary Distribution Program, NIDDK). For the radioimmunoassays, data are given as mean $\pm$ s.e.m. Serum from at least 5 mice in each category was used in the comparisons.



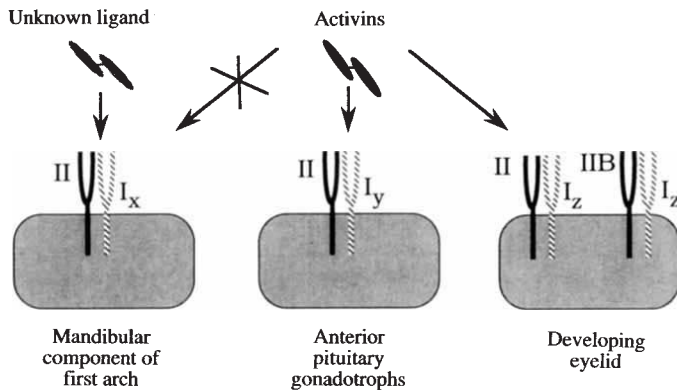


FIG. 4 Signalling through ActRcII and ActRcIIB in the developing mouse. Based on the known distribution of ActRcII and ActRcIIB^{7,9,14,16} and the results described in this and the accompanying Letter⁸, a model is postulated for signal transduction through ActRcII and ActRcIIB in combination with different type I serine/threonine kinase receptors, indicated as I_x , I_y and I_z . Activins would signal through an ActRcII (II)/type I_y complex in the anterior pituitary gonadotrophs and through both ActRcII/type I_z and ActRcIIB (IIB)/type I_z complexes in the developing eyelid. In the mandible, ActRcII/type I_x is the signalling complex. An unknown TGF- β superfamily member ligand signals through the ActRcII/type I_x receptor complex in the mandible; known activins are unable to signal through this complex. As the tissue distribution of type I receptors is not known, the type I_x , I_y and I_z receptors cannot be identified.

and incisors, and have cleft palates). However, *actRcII*^{ml}/*actRcII*^{ml} male and female mice were viable and overtly normal, although they were under-represented at weaning on both hybrid (C57/129) and 129Sv-inbred genetic backgrounds in multiple intercrosses from three independent ES cell lines (Fig. 1b).

To determine why *actRcII*^{ml}/*actRcII*^{ml} mice were under-represented, caesarian sections were done at E18.5 on *actRcII*^{ml}/+ females mated with *actRcII*^{ml}/*actRcII*^{ml} males. Of 73 offspring, 37 were *actRcII*^{ml}/*actRcII*^{ml} (51%), and eight of these (22%) showed variable hypoplasia of the mandible (micrognathia; Fig. 2) and other skeletal and facial abnormalities (cleft palate and eyelid closure defects) and died within minutes of delivery. Mice with micrognathia had secondary defects, including cleft palates and absence of incisors. Other abnormalities included defects in Meckel's cartilage and, in four of these eight mice, eyelid closure defects. These malformations are reminiscent of the Pierre-Robin syndrome in humans^{5,6}, which is associated with deficiencies in the first branchial arch. As ActRcII is expressed in the mandibular component of the first branchial arch at embryonic day 10.5 (ref. 7, and A.J.M. van den Eijnden-van Raaij, personal communication), the skeletal abnormalities, including defects in Meckel's cartilage, may be associated with altered signal transduction through ActRcII during early mandible development. These deficiencies are distinct from those seen in activin ligand-deficient mice⁸ which have primary defects in incisors, palate, eyelids and whiskers (known sites of activin expression^{7,9}). None of the activin ligand-deficient mice showed any mandible defects.

The anterior pituitary and reproductive tracts are responsive to activin^{4,10-13}, and both tissues express ActRcII and ActRcIIB^{7,9,14-16}. To identify the role of ActRcII in the activin signalling pathway in reproduction, male and female *actRcII*^{ml}/*actRcII*^{ml} mice were investigated. Male mice were delayed in reaching fertility: whereas the mean age of fertility (fertile copulations) for *actRcII*^{ml}/+ males was 59 days (range, 48–78 days; $n=7$), the mean age for the *actRcII*^{ml}/*actRcII*^{ml} males was 78 days (range, 54–130 days; $n=6$); *actRcII*^{ml}/*actRcII*^{ml} males had smaller testes from 21 days until the adult stage (Fig. 3a, b) but stages of spermatogenesis were essentially normal compared to controls (Fig. 3c, d). The seminiferous tubule diameter and total volume were reduced, consistent with an overall decrease in Sertoli cells (mitotically active before p15; ref. 17). As the epididymides from 42-day-old *actRcII*^{ml}/*actRcII*^{ml} male mice were usually devoid of spermatozoa, the delay in fertility is most likely to be due to a reduction in the seminiferous tubule volume, leading to a decrease in the amount of spermatozoa.

In contrast to *actRcII*^{ml}/*actRcII*^{ml} males, *actRcII*^{ml}/*actRcII*^{ml} females are infertile; *actRcII*^{ml}/*actRcII*^{ml} females had thin uteri and small ovaries compared to controls (Fig. 3e–h). *actRcII*^{ml}/*actRcII*^{ml} female mice had normal oocytes and essentially normal follicular development, except that follicular atresia was often seen (Fig. 3h), and the ovaries rarely had corpora lutea, confirming that *actRcII*^{ml}/*actRcII*^{ml} females did not undergo normal oestrous cycles.

Activins stimulate pituitary synthesis of follicle-stimulating hormone (FSH)⁴, and therefore the small testes and ovarian follicular atresia in *actRcII*^{ml}/*actRcII*^{ml} mice could be due to decreased FSH levels¹⁷. Although levels of luteinizing hormone were normal in the pituitary gonadotrophs of wild-type males (Fig. 3i) and *actRcII*^{ml}/*actRcII*^{ml} mice (Fig. 3j), FSH was suppressed in *actRcII*^{ml}/*actRcII*^{ml} mice (Fig. 3l) compared to controls (Fig. 3k). Serum FSH was suppressed in both female (35.2 ± 5.2 ng ml⁻¹) and male (26.8 ± 8.0 ng ml⁻¹) adult *actRcII*^{ml}/*actRcII*^{ml} mice compared to wild-type female (83.4 ± 20.6 ng ml⁻¹) and male (78.0 ± 12.0 ng ml⁻¹) controls. Our results also show that spermatogenesis can occur despite reduced FSH and without signal transduction through ActRcII.

The reproductive defects in our ActRcII mice are consistent with expression of this receptor and the role of activin in stimulating FSH synthesis^{4,7,9-16}. But the embryonic defects, manifested in ~20% of ActRcII-deficient embryos, are not consistent with defects found in activin-deficient mice (accompanying Letter), which we expected to phenocopy ActRcII-deficient mice. In particular, ActRcII-deficient mice have primary alterations in mandible development which cause secondary effects, such as a cleft palate and lack of incisors, whereas activin-deficient mice have primary defects in palate and incisor development and no mandible defects (ref. 18, and accompanying Letter), consistent with the expression pattern of activins^{7,9}. This apparent contradiction of activins signalling through ActRcII in the gonadotrophs of the pituitary but not during embryonic development, indicates that there could be different receptor complexes which respond to activins and that an unknown ligand is interacting with ActRcII during mandible development. Figure 4 presents some of the established and proposed receptor/ligand interactions in the pituitary and during embryonic development. The known requirement of the type-I receptor for signal transduction, the existence of at least two type-I receptors which bind activin, and the presence of a second type-II receptor, ActRcIIB¹⁻³, all reinforce the idea that there are several candidate proteins whose functions can be genetically tested in these signalling pathways during embryonic development. Although overexpression of a truncated ActRcII in *Xenopus laevis* inhibits mesoderm induction and promotes neural differentiation^{19,20} ActRcII-deficient mice do not have primary defects in mesoderm formation or neural development. This suggests that the overexpressed truncated ActRcII might cause phenotypic effects by interfering with signalling from other TGF- β superfamily members²¹. □

Received 9 November 1994; accepted 24 January 1995.

- Mathews, L. S. *Endocr. Rev.* **15**, 310–325 (1994).
- ten Dijke, P. et al. *Science* **264**, 101–104 (1994).
- Carcamo, J. et al. *Molec. cell. Biol.* **14**, 3810–3821 (1994).
- Vale, W., Hsueh, A., Rivier, C. & Yu, J. in *Peptide Growth Factors and their Receptors II* (eds Sporn, M. B. & Roberts, A. B.) 211–248 (Springer, Berlin, 1990).
- Caouette-Laberge, L., Bayet, B. & Laroque, Y. *Plastic Reconstr. Surg.* **93**, 934–1042 (1994).
- Clarke, L., Hepworth, W. B., Carey, J. C. & Seegmiller, R. E. *Teratology* **38**, 565–570 (1988).
- Feijen, A., Goumans, M. J. & van den Eijnden-Van Raaij, A. J. M. *Development* **120**, 3621–3637 (1994).

8. Matzuk, M. M. et al. *Nature* **374**, 354–356 (1995).
9. Roberts, V. J. & Barth, S. *Endocrinology* **134**, 914–923 (1994).
10. Woodruff, T. K., Lyon, R. J., Hansen, S. E., Rice, G. C. & Mather, J. P. *Endocrinology* **127**, 3196–3205 (1990).
11. Mather, J. P. et al. *Endocrinology* **127**, 3206–3214 (1990).
12. Kaipia, A., Toppari, J., Huhtaniemi, I. & Pranko, J. *Endocrinology* **134**, 2165–2170 (1994).
13. Shikone, T. et al. *Molec. Endocr.* **8**, 983–995 (1994).
14. deWinter, J. P. et al. *Molec. Cell Endocr.* **83**, R1–R8 (1992).
15. Feng, Z.-M., Madigan, M. B. & Chen, C.-L. *Endocrinology* **132**, 2593–2600 (1993).
16. Cameron, V. A. et al. *Endocrinology* **134**, 799–808 (1994).
17. Ritzen, E. M., Hansson, V. & French, F. S. in *The Testis* 2nd edn (eds Burger, H. & de Kretser, D.) 269–302 (Raven, New York, 1989).
18. Vassalli, A. et al. *Genes Dev.* **8**, 414–427 (1994).
19. Hemmati-Brivanlou, A. & Melton, D. A. *Nature* **359**, 609–614 (1992).
20. Hemmati-Brivanlou, A. & Melton, D. A. *Cell* **77**, 273–281 (1994).
21. Schulte-Merker, S., Smith, J. C. & Dale, L. *EMBO J.* **13**, 3533–3541 (1994).
22. Matzuk, M. M. & Bradley, A. *Biochem. biophys. Res. Commun.* **185**, 404–413 (1992).
23. Matzuk, M. M., Finegold, M. J., Su, J.-G. J., Hsueh, A. J. W. & Bradley, A. *Nature* **360**, 313–319 (1992).
24. Mathews, L. S. & Vale, W. W. *Cell* **65**, 973–982 (1991).
25. Bradley, A. in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach* (ed. Robinson, E. J.) 113–151 (IRL, Oxford, 1987).
26. Ramirez-Solis, R. et al. *Analyt. Biochem.* **201**, 331–335 (1992).
27. Ramirez-Solis, R. et al. *Cell* **83**, 279–294 (1993).
28. Kumar, T. R., Fairchild-Huntress, V. & Low, M. *Molec. Endocr.* **6**, 81–90 (1992).

ACKNOWLEDGEMENTS. We thank J. van den Eijnden-van Raaij for sharing activin-receptor expression data and other information before publication; N. Lu and R. Towns for technical assistance; C. McDonnell for help with cryostat sectioning; S. Topouzis for help with northern blot analysis; S. Baker and B. Powell for help with manuscript preparation; and R. Behringer and M. Meistrich for advice and review of the manuscript. FSH radioimmunoassay reagents were a gift from the National Hormone and Pituitary Distribution Program, National Institute of Diabetes and Digestive and Kidney Diseases. This research was supported in part by grants (to M.M.M.) from the NIH and the Lalor Foundation. A.B. is an investigator with the Howard Hughes Medical Institute.

Multiple defects and perinatal death in mice deficient in follistatin

Martin M. Matzuk^{*†‡}, Naifang Lu[†], Hannes Vogel[†], Klaus Sellheyer[†], Dennis R. Roop[†] & Allan Bradley^{*§}

Departments of ^{*} Molecular and Human Genetics, [†] Pathology, [‡] Cell Biology, [§] Howard Hughes Medical Institute, Baylor College of Medicine, Houston, Texas 77030, USA

FOLLISTATIN, an activin-binding protein and activin antagonist *in vitro*^{1,2}, can bind to heparan sulphate proteoglycans³ and may function *in vivo* to present activins to their receptors. In the mouse, follistatin messenger RNA is first detected in the deciduum (on embryonic day 5.5), and later in the developing hindbrain, somites, vibrissae, teeth, epidermis and muscle^{4–11}. In *Xenopus laevis*, overexpression of follistatin leads to induction of neural tissue¹². Here we use loss-of-function mutant mice to investigate the function of follistatin in mammals. We find that follistatin-deficient mice are retarded in their growth, have decreased mass of the diaphragm and intercostal muscles, shiny taut skin, skeletal defects of the hard palate and the thirteenth pair of ribs, their whisker and tooth development is abnormal, they fail to breathe, and die within hours of birth. These defects are more widespread than those seen in activin-deficient mutant mice, indicating that follistatin may modulate the actions of several members of the transforming growth factor- β family.

To define the roles of follistatin in mammalian development, a targeted deletion (fs^{ml}) of the 6-exon follistatin gene was generated using embryonic stem (ES) cell technology (Fig. 1a), and mice heterozygous for this deletion ($fs^{ml}/+$) were intercrossed to generate mice homozygous for the deleted follistatin allele (fs^{ml}/fs^{ml}). Genotyping of mice at birth and at embryonic day 18.5 (E18.5) revealed that fs^{ml}/fs^{ml} mice (follistatin-deficient) could survive to birth but died within hours of delivery (Fig. 1b). Of 196 pups at E18.5, 44 were homozygotes (22.4%), 97 were heterozygote (49.5%) and 55 were wild-type (28.1%), consistent with the expected mendelian frequency of 1:2:1. Low-

FIG. 1 Targeted deletion of the follistatin gene in ES cells and Southern blot analysis of DNA from offspring derived from heterozygous matings. a, The targeting vector contains 3.8 kb of isogenic DNA homologous to the 5' non-translated sequence of the mouse follistatin gene, 3.4 kb of sequence homologous to the 3' end of the mouse follistatin gene, a PGK-*hprt* (where PGK represents phosphoglycerate kinase I promoter) expression cassette, and an MC1-tk (thymidine kinase) expression cassette. Homologous recombination between this targeting vector and the endogenous follistatin gene in mouse ES cells should result in a 5.1-kb deletion of the 6 exons coding for follistatin, ensuring that no follistatin mRNA and therefore no follistatin protein was produced in animals homozygous for the targeted allele. Introduction of new diagnostic restriction endonuclease sites (*Bgl*II and *Xba*I) were used to differentiate wild-type versus recombinant alleles. Six ES cell clones out of 1,290 clones (1:215) screened by Southern blot analysis^{21,22} were correctly targeted. One ES cell clone, FS3-C2, gave rise to multiple male chimaeras and germline transmission of the deleted allele was demonstrated for 3 of 4 chimaeric males. Male and female mice heterozygous for the deleted follistatin allele occurred at the expected ratio and were viable and fertile. b, Genomic (tail) DNA (~5 μ g) from offspring from a single litter was restricted with *Xba*I and *Eco*RI and analysed by Southern blot analysis²¹ using a 3' probe as shown. The presence of a 5.9-kb fragment versus a 3.7-kb wild-type fragment is diagnostic of the deletion when using the 3' probe. Homozygote offspring are dead within hours of birth. WT, wild type; +/–, heterozygote; –/–, homozygote mutant. c, Southern blots of wild-type (WT) and fs^{ml}/fs^{ml} (–/–) mice, chicken and *Xenopus laevis* genomic DNA (5 μ g per lane) digested with *Bam*HI were analysed using a human follistatin cDNA. The absence of any hybridizable fragment bands in the lane genotyped as homozygote (–/–) confirmed that these mice lack the 6 exons coding for the follistatin gene, that this is a null allele (fs^{ml}), and that there were no follistatin-related genes in the mouse.

METHODS. More than 20 kb of DNA encompassing the 6-exon mouse follistatin gene sequences was isolated from a 129SvEv genomic library using a follistatin cDNA. Linearized vector (25 μ g) was electroporated into the *hprt*-negative AB2.1 ES cell line, selected in HAT and FIAU (where HAT represents hypoxanthine, aminopterin, thymidine, and FIAU is (1-(2'-deoxy-2'-fluoro- β -D-arabinofuranosyl)-5-iodouracil), and clones were injected into blastocysts to generate chimaeras²¹. Enrichment in HAT and FIAU was 32-fold compared to HAT alone. Southern blot analysis of the ES cell clones was as described^{21,22}. Low-stringency hybridization of the mouse, chicken and *Xenopus laevis* blot was by overnight hybridization at 60 °C followed by 4 washes for 30' each at 50 °C with a 2 $\times$ SSC, 1% SDS solution.

stringency hybridization with a follistatin complementary DNA could detect genomic fragments in wild-type mouse, chicken and *Xenopus laevis* but not in the fs^{ml}/fs^{ml} mice, confirming that we had deleted the only follistatin gene present in the mouse (Fig. 1c).

fs^{ml}/fs^{ml} newborns could be phenotypically scored because they were growth-retarded and had shiny, taut skin similar to that seen in the human condition of restrictive dermopathy¹³ (Fig. 2a). The phenotypic effects of follistatin deficiency were independent of the genetic background (that is, the phenotype was similar on 129SvEv inbred, C57/129 hybrid and C57B1/6 (3 backcrosses) genetic backgrounds). At birth, most fs^{ml}/fs^{ml} mice remained pale and cyanotic, the lungs of the fs^{ml}/fs^{ml} mice sank in liquid and, on histological examination, the alveolar spaces were poorly expanded, consistent with poor breathing, although primary pulmonary defects could not be detected (data not shown). Analysis of the central and peripheral nervous system by gross dissection and histological procedures did not detect any significant abnormalities. Immunohistochemistry of E10.5 whole-mount embryos using an antibody raised against the 155K neurofilament protein¹⁴ indicated that cranial nerves V, VII, IX and X and the spinal ganglia were developing normally (data not shown).

Defects were detected in the musculoskeletal system. fs^{ml}/fs^{ml} mice lacked incisors (6 of 34, hybrid background; Fig. 2b) or had delayed incisor development (Table 1). Three of 19 (16%) hybrid background fs^{ml}/fs^{ml} mice had a cleft secondary palate. Six out of 11 129SvEv inbred mice (55%) and 6 out of 28 C57/

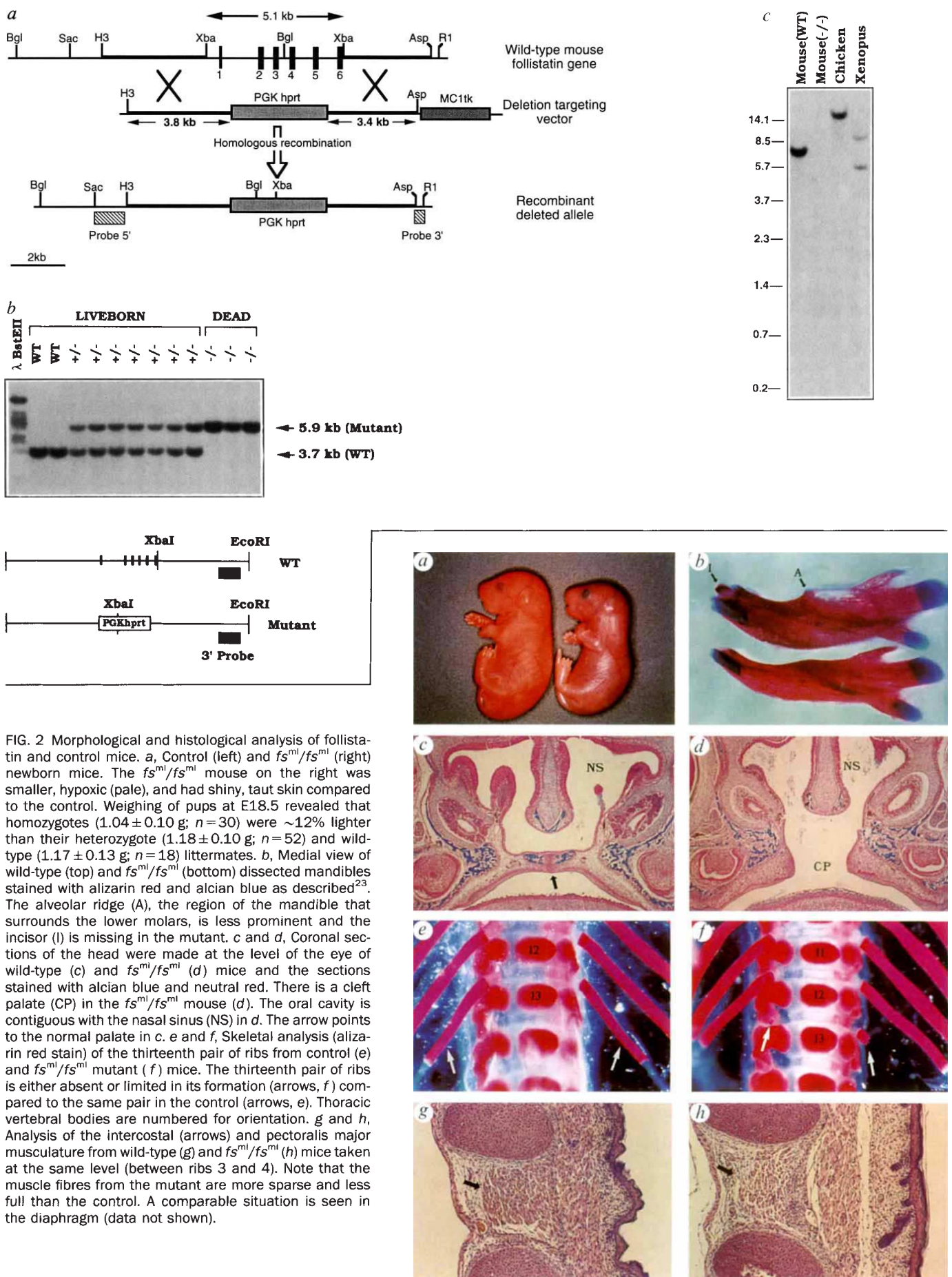


FIG. 2 Morphological and histological analysis of follistatin and control mice. **a**, Control (left) and fs^{mi}/fs^{mi} (right) newborn mice. The fs^{mi}/fs^{mi} mouse on the right was smaller, hypoxic (pale), and had shiny, taut skin compared to the control. Weighing of pups at E18.5 revealed that homozygotes (1.04 ± 0.10 g; $n = 30$) were $\sim 12\%$ lighter than their heterozygote (1.18 ± 0.10 g; $n = 52$) and wild-type (1.17 ± 0.13 g; $n = 18$) littermates. **b**, Medial view of wild-type (top) and fs^{mi}/fs^{mi} (bottom) dissected mandibles stained with alizarin red and alcian blue as described²³. The alveolar ridge (A), the region of the mandible that surrounds the lower molars, is less prominent and the incisor (I) is missing in the mutant. **c** and **d**, Coronal sections of the head were made at the level of the eye of wild-type (**c**) and fs^{mi}/fs^{mi} (**d**) mice and the sections stained with alcian blue and neutral red. There is a cleft palate (CP) in the fs^{mi}/fs^{mi} mouse (**d**). The oral cavity is contiguous with the nasal sinus (NS) in **d**. The arrow points to the normal palate in **c**. **e** and **f**, Skeletal analysis (alizarin red stain) of the thirteenth pair of ribs from control (**e**) and fs^{mi}/fs^{mi} mutant (**f**) mice. The thirteenth pair of ribs is either absent or limited in its formation (arrows, **f**) compared to the same pair in the control (arrows, **e**). Thoracic vertebral bodies are numbered for orientation. **g** and **h**, Analysis of the intercostal (arrows) and pectoralis major musculature from wild-type (**g**) and fs^{mi}/fs^{mi} (**h**) mice taken at the same level (between ribs 3 and 4). Note that the muscle fibres from the mutant are more sparse and less full than the control. A comparable situation is seen in the diaphragm (data not shown).

129 hybrid genetic background mice (21%) lacked a hard palate similar to activin- β A-deficient mice¹⁵ (see below; Fig. 2c, d). Both inbred and hybrid fs^{ml}/fs^{ml} mice had defects in the thirteenth pair of ribs and a decrease in the number of lumbar vertebrae (Fig. 2e, f; Table 1). The size of the ribs and penetrance of this phenotype was dependent on the genetic background (Table 1). The intercostal (Fig. 2g, h) and diaphragmatic (data not shown) muscles showed a decrease in muscle mass consistent with expression of follistatin in the muscle at birth¹⁰. Histological and electron microscope analysis of the musculature, including analysis of ATPase activity, glycogen content and mitochondria, failed to demonstrate any primary defect (data not shown). Thus, it is not clear whether the decrease in muscle mass in the fs^{ml}/fs^{ml} mice is primary or secondary to an overall growth deficiency.

Whisker development was abnormal: the whiskers were too thin and were inappropriately oriented, suggesting that follistatin may be an important modulator of activin- β A action (Fig. 3a–d; see accompanying Letter). The skin showed hyperkeratosis, as indicated by thickened granular and stratum corneum layers (Fig. 3e, f). Electron microscope examination revealed a 25% increase in the stratum corneum cells compared to controls. The shiny, taut skin of fs^{ml}/fs^{ml} mice (Fig. 2a) resembles the skin of transgenic mice with directed overexpression of transforming growth factor (TGF)- β 1 to the epidermis using a human keratin-1 promoter (HK1.TGF- β 1)¹⁶. Keratin-6 expression was abnormal in the interfollicular epidermis of the fs^{ml}/fs^{ml} mice, as it was in HK1.TGF- β 1 mice (Fig. 3h). Normally

keratin-6 is found only in the outer root sheath of hair follicles (Fig. 3g). Although abnormal keratin-6 expression is often associated with hyperproliferation¹⁷, the fs^{ml}/fs^{ml} mice showed normal epidermal mitotic activity as judged by 5-bromo-deoxyuridine labelling. These results, together with those obtained for the growth-arrested epidermis of HK1.TGF- β 1 mice¹⁶, suggest that keratin-6 expression can occur in response to a variety of stimuli that perturb normal epidermal development.

In conclusion, fs^{ml}/fs^{ml} mice have defects in the musculoskeletal system and in the epidermis, and this constriction effect probably contributes to their rapid demise. Follistatin is expressed in the rhombomeres in the mid-gestational mouse⁵, and can cause differentiation of neural cell lines¹⁸; overexpression in *Xenopus laevis* suggests that follistatin may be essential for neural induction¹². Our results show that the absence of follistatin in the mouse *in vivo* does not apparently affect the gross development of the nervous system. Interestingly, the fs^{ml}/fs^{ml} mice demonstrate abnormal whisker and tooth development and hard-palate defects similar to activin- β A-deficient mice¹⁵. In the vibrissae, teeth and palate, where follistatin is expressed adjacent to activin- β A⁷, follistatin appears to play a role in activin signal transduction, possibly sequestering activins in heparan sulphate proteoglycans³ and presenting activin to activin receptors similar to type III TGF- β receptors¹⁹. The follistatin-deficient mice also have other defects not seen in activin-deficient mice¹⁵. Some of these defects are similar to those of the BMP-5 (*short ear*) mutant mice, which have defects in the thirteenth pair of ribs²⁰, and the TGF- β overexpressors, which have shiny, taut skin¹⁶. These

FIG. 3 a and b, Gross analysis of the whisker pads and face of control (a) and fs^{ml}/fs^{ml} (b) newborn mice demonstrating the disoriented whiskers and shiny skin of the mutant (b). c and d, Histological analysis (haematoxylin and eosin stain of transverse sections) of the whisker follicles of control (c) and fs^{ml}/fs^{ml} (d) newborn mice. The whisker shafts (arrow) are perpendicular to the surface in the control (c). In the mutant (d), the shaft (arrow) projects parallel before turning perpendicularly. e and f, Histology of the skin of control (e) and fs^{ml}/fs^{ml} (f) newborn mice photographed at the same magnification. Note the thickened granular (between large white arrows, left) and stratum corneum (between small white arrows, right) layers in the mutant (f) compared to the control (e). g and h, Double-label immunofluorescence of the skin from the back of E18.5 control (g) and fs^{ml}/fs^{ml} (h) mice. The antibody raised against keratin 14 (red) will stain both interfollicular epidermis and hair follicle, whereas the antibody against keratin 6 (yellow/green) will normally stain only the outer root sheath (arrow) of the hair follicle (g). In the mutant (h), abnormal keratin 6 expression seen in the interfollicular epidermis (E). Single-label immunofluorescence of the mutant skin did not reveal any abnormality in keratin 14 expression (data not shown).

METHODS. Histological analysis of the muscle and skeleton, and double-label immunofluorescence of the skin were performed as described^{16,21,23}.

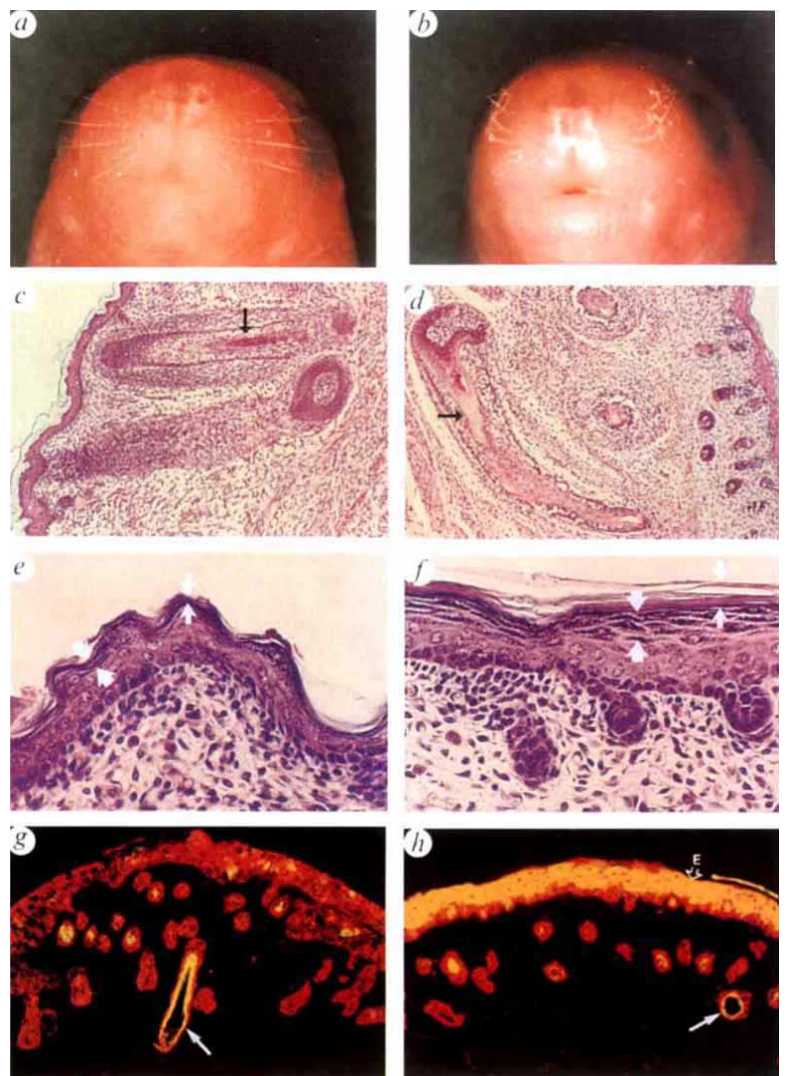


TABLE 1 Skeletal, palate and tooth defects

	Intact	13th pair of ribs		Five lumbar vertebrae*	Palate defects		Lower incisors	
		Anlage	Zero ribs		Cleft	Hard palate	Absent	Delayed
C57/129 hybrid								
Wild-type	15	0	0	2/15	0	0	0	0
Heterozygote	29	6	0	21/35	0	0	0	0
Homozygote†	3	17	8	28/28	3/19	6/28	6/34	23/34
129 inbred								
Wild-type	9	0	0	0/9		0	0	0
Heterozygote	11	4	0	13/15		0	0	0
Homozygote	0	9	2	11/11		6/11	0/11	11/11

* Six lumbar vertebrae are the normal complement in the mouse.

† In 4/28 homozygotes, one or both of the seventh ribs failed to fuse to the sternum.

similarities suggest that follistatin may be a modulator of other TGF- β -related proteins or may function independently. Thus, the results obtained by overexpression of follistatin in *Xenopus laevis*¹² may be a consequence of up or downregulation of the actions of several TGF- β -related proteins. It will be critical to compare mice deficient in other TGF- β superfamily members to our follistatin-deficient mice in order to establish the extent to which follistatin is involved in signal transduction of other TGF- β -related proteins. □

Received 9 November 1994; accepted 24 January 1995.

1. Michel, V., Farnworth, P. & Findlay, J. K. *Molec. cell. Endocr.* **91**, 1–11 (1993).
2. Nakamura, T. et al. *Science* **247**, 836–838 (1990).
3. Nakamura, T., Sugino, K., Titani, K. & Sugino, H. *J. biol. Chem.* **266**, 19432–19437 (1991).
4. van den Eijnden-van Raaij, A. J. M., Feijen, A., Lawson, K. A. & Mummery, C. L. *Dev. Biol.* **154**, 356–365 (1992).
5. Albano, R. M., Arkell, R., Beddington, R. S. P. & Smith, J. C. *Development* **120**, 803–813 (1994).
6. Feijen, A., Goumans, M. J. & van den Eijnden-van Raaij, A. J. M. *Development* **120**, 3621–3637 (1994).

7. Roberts, V. J. & Barth, S. *Endocrinology* **134**, 914–923 (1994).
8. Shimasaki, S. et al. *Mol. Endocrinology* **3**, 652–659 (1989).
9. Michel, V., Albiston, A. & Findlay, J. K. *Biochem. biophys. Res. Commun.* **173**, 401–407 (1990).
10. Michel, V., Rao, A. & Findlay, J. K. *Biochem. biophys. Res. Commun.* **180**, 223–230 (1991).
11. Tashiro, K. et al. *Biochem. Biophys. Res. Commun.* **174**, 1022–1027 (1991).
12. Hemmati-Brivanlou, A., Kelly, O. G. & Melton, D. A. *Cell* **77**, 283–295 (1994).
13. Pierard-Franchimont, C. et al. *J. Path.* **167**, 2223–2228 (1992).
14. Chisaka, O., Musci, T. S. & Capecchi, M. R. *Nature* **335**, 516–520 (1992).
15. Matzuk, M. M. et al. *Nature* **374**, 354–356 (1995).
16. Sellheyer, K. et al. *Proc. natn. Acad. Sci. U.S.A.* **90**, 5237–5241 (1993).
17. Weiss, R. A., Eichner, R. & Sun, T.-T. *J. Cell Biol.* **98**, 1397–1406 (1984).
18. Hashimoto, M. et al. *J. biol. Chem.* **267**, 7203–7206 (1992).
19. Lopez-Casillas, F., Wrana, J. L. & Massague, J. *Cell* **73**, 1435–1444 (1993).
20. Kingsley, D. M. et al. *Cell* **71**, 399–410 (1992).
21. Matzuk, M. M., Finegold, M. J., Su, J.-G. J., Hsueh, A. J. W. & Bradley, A. *Nature* **360**, 313–319 (1992).
22. Ramirez-Solis, R. et al. *Analyt. Biochem.* **201**, 331–335 (1992).
23. Ramirez-Solis, R. et al. *Cell* **83**, 279–294 (1993).

ACKNOWLEDGEMENTS. We thank S. Baker and B. Powell for assistance with manuscript preparation; W. Moyle and R. Myers for their help in isolating follistatin cDNA; R. Behringer, J. Bickenbach and T. Rajendra Kumar for critically reviewing the manuscript; and J. Bickenbach for advice on histology and electron microscopy. This research was supported in part by grants from the NIH (to M.M.M.). A.B. is an associate investigator with the Howard Hughes Medical Institute.

The product of *hedgehog* autoproteolytic cleavage active in local and long-range signalling

Jeffery A. Porter, Doris P. von Kessler, Stephen C. Ekker, Keith E. Young, John J. Lee, Kevin Moses* & Philip A. Beachy

Howard Hughes Medical Institute, Department of Molecular Biology and Genetics, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA

* Department of Biological Sciences, University of Southern California, Los Angeles, California 90089, USA

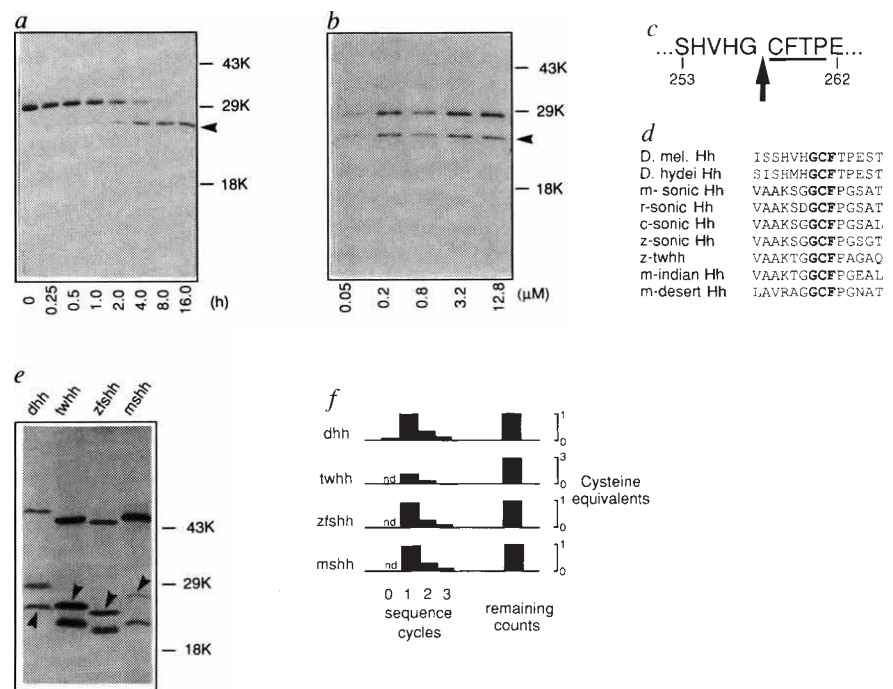
THE secreted protein products of the *hedgehog* (*hh*) gene family are associated with local and long-range signalling activities that are responsible for developmental patterning in multiple systems, including *Drosophila* embryonic and larval tissues^{1–8} and vertebrate neural tube, limbs and somites^{9–15}. In a process that is critical for full biological activity, the *hedgehog* protein (Hh) undergoes autoproteolysis to generate two biochemically distinct products, an 18K amino-terminal fragment, N, and a 25K carboxy-terminal fragment, C (ref. 16); mutations that block autoproteolysis impair Hh function. We have identified the site of autoproteolytic cleavage and find that it is broadly conserved throughout the *hedgehog* family. Knowing the site of cleavage, we were able to test the

function of the N and C cleavage products in *Drosophila* assays. We show here that the N product is the active species in both local and long-range signalling. Consistent with this, all twelve mapped *hedgehog* mutations either affected the structure of the N product directly or otherwise blocked the release of N from the Hh precursor as a result of deletion or alteration of sequences in the C domain.

To characterize the signalling molecules produced from the uncleaved Hh precursor U, we determined the site of autoproteolysis and the structures of the cleaved products using a purified Hh protein from a bacterial source. This purified protein, His₆C, contains primarily carboxy-terminal sequences and can generate a 25K cleavage product corresponding to the native C cleavage product produced in *Drosophila* cells (Fig. 1a). In reactions lasting 3 h and using a wide range of concentrations of starting material, this protein displayed kinetics that were independent of concentration, strongly suggesting that the cleavage was intramolecular (Fig. 1b).

Amino-terminal sequencing of the 25K cleavage product showed that Hh is cut between Gly 257 and Cys 258 (Fig. 1c). Sequence alignment with other insect and vertebrate *hh* proteins demonstrates the absolute conservation of the Gly-Cys-Phe sequence at the site of autoproteolysis (Fig. 1d). To test for the presence of a cysteine residue at the amino terminus of C fragments derived from *Drosophila* (*hh*), zebrafish (*twhh* and *shh*) (S.C.E., manuscript submitted) and mouse (*shh*) genes, ³⁵S-cysteine was incorporated during *in vitro* translation (Fig. 1e). Figure 1f shows that the first round of protein sequencing in each case released the proportion of total incorporated ³⁵S corresponding to a single cysteine residue, establishing that cysteine

FIG. 1 Characterization of *Drosophila* and vertebrate Hh autoproteolytic cleavage. **a**, Coomassie-blue-stained gel showing time course of autoproteolytic cleavage of His₆C, a bacterially expressed derivative of *Drosophila* hh protein carrying predominantly C-terminal sequences with an N-terminal fusion of a His₆ purification tag. **b**, Coomassie-blue-stained gel used to monitor the extent of His₆C autoproteolytic cleavage after a 3-h incubation with starting concentrations spanning a 256-fold range. The concentration-independent kinetics are consistent with an intramolecular mechanism of cleavage. **c**, Site of autoproteolytic cleavage in the His₆C protein as determined by automated N-terminal sequencing of the 25K cleavage product. The residues recovered during sequence determination are underlined and the arrow denotes the cleavage site. **d**, Alignment of known *hedgehog* protein sequences, showing absolute conservation of the tripeptide Gly-Cys-Phe at a location corresponding to the *Drosophila* cleavage site. Abbreviations are as follows: D. mel., *Drosophila melanogaster*; D. hydei., *Drosophila hydei*; m., mouse; r., rat; c., chicken; z., zebrafish; z-twihh, zebrafish *tiggy-winkle* hh. **e**, Autoprocessing of four hh proteins translated *in vitro* in the presence of ³⁵S-cysteine. C fragments are denoted by arrowheads. Whereas in the absence of signal cleavage the N-terminal fragment of *Drosophila* Hh migrates with reduced mobility relative to C (ref. 16), C fragments from the three vertebrate proteins are larger than N fragments, as previously shown^{13,16}. **f**, Automated N-terminal sequencing of the ³⁵S-labelled C fragments generated in **e**. A peak of radioactivity is released in the first round of degradation for each of the four proteins, indicating the presence of ³⁵S-cysteine at the N terminus of each C fragment. Counts removed in the first cycle are expressed as cysteine equivalents, where 1 equivalent equals the total counts divided by the number of predicted cysteines present (dhh has 2; twhh has 4).



METHODS. Bacterial Hh-fusion protein was obtained using a pRSET B vector (Invitrogen) containing hh sequences encoding amino acids 83–471 with residues 89–254 deleted. The soluble, bacterially expressed protein (M_r 29K) was purified to homogeneity by chromatography with immobilized Ni²⁺ resin and mono-S. For time course and concentration dependence experiments, DTT was added (50 mM final concentration) to the fusion protein at 23 °C to initiate the reactions. At various times (a) or at 3 h (b) samples were combined with SDS-sample buffer to terminate the reaction (dilute samples from b were first concentrated with trichloroacetic acid to give roughly equal loading). For automated N-terminal sequencing, overnight cleavage reactions were electrophoresed on 14% SDS-polyacrylamide gel and transferred to PVDF membrane. The 25K C fragment was stained with Ponceau S, excised and subjected to 4 rounds of Edman degradation in an automated protein sequencer (Applied Biosystems). For *in vitro* translation ³⁵S-cysteine was used¹⁶. For automated N-terminal sequencing, the C fragments from these reactions were processed as described. A pre-cycle, which included solvent washes of the PVDF membrane in the absence of degradation, was performed with *Drosophila* cleavage product C as a control to assay for counts removed irrespective of sequence.

is also the first residue of fragment C in the vertebrate proteins. This limited sequence information, together with the apparent sizes of the cleaved vertebrate fragments, pinpoints the site of autoproteolytic cleavage to the bond preceding the cysteine residue in the conserved Gly-Cys-Phe sequence.

To ascertain the requirements for Hh protein *in vivo*, constructs designed to produce the N cleavage product in the absence of C, the C cleavage product in the absence of N, and a precursor rendered inert by replacement of Cys at the cleavage site (U_{CA}; Fig. 2a) were tested in translation assays (Fig. 2b) and by immunoblotting heat-shock-induced proteins from transgenic *Drosophila* embryos (Fig. 2c, d). Transgenic embryos carrying the heat-shock-inducible constructs were tested for the following phenotypic effects: (1) expansion of *wingless* expression⁵ (Fig. 3a–e), a local signalling activity; (2) reprogramming of cell fates in the dorsal cuticle⁷ (Fig. 3k–o), which involves both local and long-range activities; (3) reprogramming of ventral cuticle fates⁵ (Fig. 3f–j); and (4) changes in imaginal disc morphology and gene expression⁸ (data not shown). In each of these assays, individuals ectopically expressing fragment N produced phenotypes identical to those reported for individuals ectopically expressing wild-type Hh. In contrast, the C and U_{CA} constructs did not produce these effects, suggesting that N is both necessary and sufficient for local and long-range signalling in *Drosophila*.

In support of these conclusions, we found that the primary sequence lesions in a collection of twelve hh mutant alleles^{3,17}

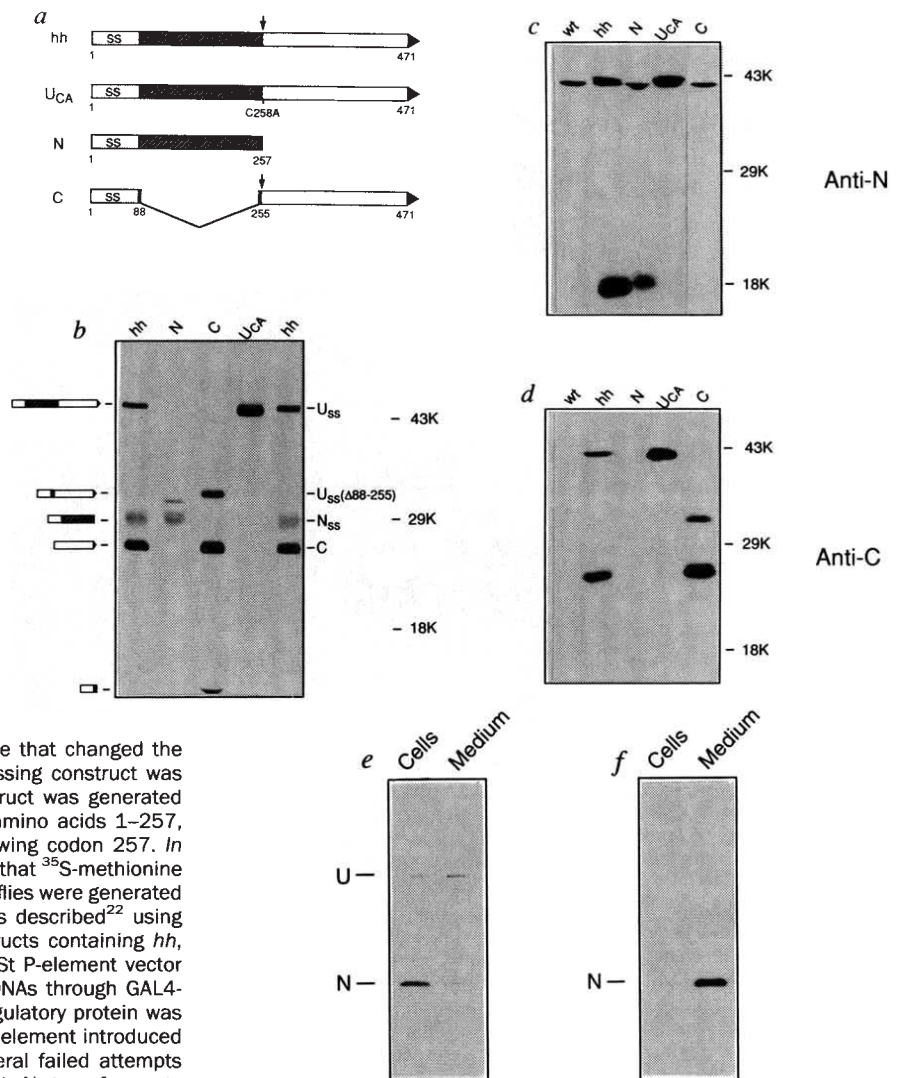
TABLE 1 Analysis of hh mutant alleles

Allele	Strength	Lesion	Protein product	Autocleavage (IVT)	Ref.
I					
9k	w	P109L		+	17
ts2	w	TNSASG115-120/IRV		+	3
302	s	E202K		+	3
278	s	A205V		+	3
287	w	F230I		+	3
20-107	m	S244F		+	17
10B	w/m	ex2/int2 splice mutant		NA	17
13C	s	W232(amber)		NA	17
279	s	W232(opal)		NA	3
II					
281	m	H329Y		—	3
068	m/s	C400Y		—	3
13E	m	Y457(ochre)		-/+	17

The primary sequence lesions in the open reading frames of previously identified hh mutant alleles^{3,17} were determined and the resulting mutant proteins tested for autoproteolytic activity *in vitro*. For each allele analysed, the table lists the strength of the mutation (w, m and s for weak, moderate and strong, respectively), the effect of the mutant lesion on amino-acid sequence, a diagram of the mutant protein, and ability of the mutant protein to undergo autoproteolytic cleavage. Allele strength is based upon severity of the ventral cuticle phenotype analysed at 22 °C. The hh^{10B} allele is unique in that it disrupts a consensus splice junction, presumably resulting in read-through into intron sequences that contain a stop codon. Sequences of the mutant alleles were derived from PCR-amplified genomic DNA clones, with a minimum of 6 independent PCR products sequenced for each allele.

FIG. 2 Expression of the N, C, U_{CA} and wild-type *hh* constructs *in vitro* and *in vivo*. **a**, Constructs derived from *Drosophila hh* designed specifically to express N, C and U_{CA}, an uncleaved form of Hh precursor. The arrow marks the wild-type autoproteolytic cleavage site and SS denotes the signal sequence. **b**, *In vitro* translation reactions with constructs in **a** give the N, C and U_{CA} products. Note that signal sequences are not removed *in vitro* in the absence of microsomes. **c** and **d**, Immunoblots using N- and C-specific antibodies with heat-shock-induced proteins extracted from transgenic flies carrying these constructs. Note that a background band migrating just beneath U is detected by anti-N antibody in **c** but not by anti-C in **d**. Endogenous Hh is too low to be detected in these experiments. **e**, **f** Cell-association of N is dependent upon autoproteolytic cleavage by C. Immunoblots using N-specific antibodies with extracts from cells or medium collected from S2 cell cultures expressing Hh (**e**) or Hh-N (**f**). N is found predominantly associated with the cells when expressed from a WT *hh* construct (**e**) but is primarily in the culture medium when synthesized from a construct that precisely deletes the C domain (**f**). 'U' denotes the small amount of uncleaved Hh precursor that is evenly distributed between fractions.

METHODS. The U_{CA}-expressing construct was generated using *hh* in pBluescript as template for PCR-mediated mutagenesis with an oligonucleotide that changed the Cys 258 codon to an alanine codon. The C-expressing construct was generated as described¹⁶; the N-expressing construct was generated by PCR amplification of *hh* sequences encoding amino acids 1–257, with a chain-termination codon incorporated following codon 257. *In vitro* translation was as described for Fig. 1, except that ³⁵S-methionine (NEN) was used instead of ³⁵S-cysteine. Transgenic flies were generated by P-element-mediated germline transformation as described²² using heat-shock-promoter expression vectors for constructs containing *hh*, C, and U_{CA} inserts²³. For the N construct, the pUAS P-element vector was used to allow for expression of subcloned cDNAs through GAL4-responsive elements²⁴. Expression of the GAL4 regulatory protein was under heat-shock-inducible control from another P-element introduced by mating. The GAL4 system was used after several failed attempts (1,500 injectees) to generate heat-shocked (hs) N transformants directly. Extracts for **c** and **d** were from embryos of wild-type Canton S, from flies homozygous (for hsHh, hsC or hsU_{CA} P-elements, and from flies containing one copy each of the UAS N and hsGAL4 P-elements. Embryo collections (0–16 h) were heat-shocked for 30 min at 37 °C, and embryos from wild-type, hsHh, hsC and hsU_{CA} were incubated at 25 °C for 1 h; the UAS N/hsGAL4 flies were incubated for 2 h to allow for secondary induction of N. Embryos were dechorionated and combined with 10 vol SDS-sample buffer. Proteins were



electrophoresed in SDS-polyacrylamide gels, transferred to nitrocellulose, incubated with N- and C-specific primary antibodies followed by peroxidase-conjugated secondary antibody, and visualized by ECL (Amersham) and autoradiography. For **e**, **f**, Hh and Hh-N coding sequences were subcloned into the pMK33 vector, stably transfected into S2 cells, and induced to express¹⁶; cell and medium samples represent equal volumes of the total culture and were generated as before¹⁶.

affected either the structure of the N cleavage product (class I) or the efficiency of N release from U by autoproteolytic cleavage (class II; Table 1). The class I lesions include two mutations that truncate N (*hh*^{13C} and *hh*²⁷⁹), and a third mutation expected to block splicing of the second to the third exons, thereby introducing missense and a truncation (*hh*^{10B}). The remaining six class I proteins undergo autoproteolytic cleavage *in vitro*, but introduce substitutions at positions in N that are invariant among all *hh* proteins. These sequence changes presumably affect the stability or folded structure of N, or perturb residues that contact a receptor. *hh*^{9K} and *hh*¹⁵² are thermosensitive and affect nearby residues in N, suggesting a proclivity of this region to thermosensitive mutation. Whereas one of the class II mutations introduces truncation of C (*hh*^{13E}), *hh*⁶⁰⁸ substitutes tyrosine for Cys 400, a residue invariant among *hh* family members but not essential for autoproteolytic cleavage (S.C.E., J.J.L. and P.A.B., unpublished data) and *hh*²⁸¹ changes His 329, a residue previously shown to be required for autoproteolytic cleavage and full biological activity of Hedgehog¹⁶, to tyrosine.

The class II mutations and U_{CA} transgenic flies support our proposal that N must be released from U for activity. As shown

previously, autoproteolytic cleavage primarily requires determinants within C, and specifically requires His 329 (ref. 16). Autocleavage of His₃₂₉C was not inhibited by any of a wide range of standard protease inhibitors, nor by 2 M urea, nor by a peptide containing the cleavage site, but it did require high concentrations of dithiothreitol (50 mM) and could be accelerated by addition of hydroxylamine (data not shown). Together with the essential role of Cys 258, these observations suggest a novel cleavage mechanism involving activation by His 329 of the Cys 258 thiol for nucleophilic attack at the carbonyl of Gly 257, reminiscent of the first step in serinolysis, the autocleavage mechanism of prohistidine decarboxylase¹⁸. Computer-assisted sequence analysis has revealed a conserved sequence motif at and around the site of *hedgehog* autocleavage (including the essential Cys 258 residue) and at the amino-terminal splice junction of self-splicing proteins¹⁹; protein splicing also may proceed through an ester intermediate at this junction²⁰. In Hh, the thioester intermediate could simply be hydrolysed, as in thiol proteolysis, or be attacked by another nucleophile, resulting in modification of N. Interestingly, whereas N synthesized from a native construct is primarily cell-associated (Fig. 2e)¹⁶, N gener-

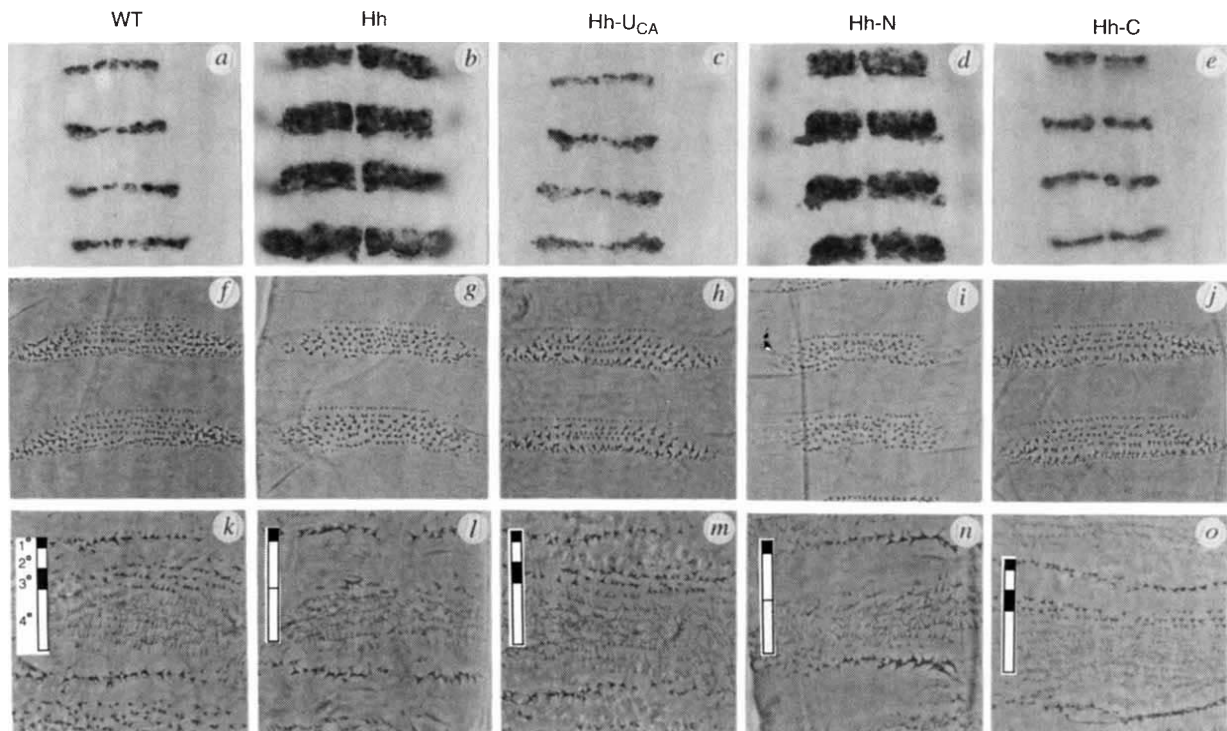


FIG. 3 Embryonic effects of ubiquitous Hh, U_{CA} , N, and C expression. Control embryos (a, f, k; labelled WT), embryos expressing Hh (b, g, l; labelled Hh), embryos expressing U_{CA} (c, h, m; labelled Hh- U_{CA}), embryos expressing N (d, i, n; labelled Hh-N), and embryos expressing C (e, j, o; labelled Hh-N) were analysed for expansion of embryonic *wingless* expression (a-e), ventral cuticle defects (f-j) and dorsal cuticle defects (k-o) after heat-shock induction of the respective transgenes during early embryogenesis. a-e, *wingless* expression was detected by *in situ* hybridization at the extended germ-band stage. Ventral views of para-segments 5-8 show a 3-4 cell expansion of *wg* expression in Hh and Hh-N-expressing embryos. f-j, A4 and A5 segments of the ventral cuticle of first instar larvae are shown with anterior up and dorsal down. The normal trapezoidal geometry of the denticle belts were replaced by a more rectangular shape in Hh (g) and Hh-N (i); in addition, the orientation and size of denticle rows 4 and 5 in these embryos were altered

ated from a truncated construct in cultured cells predominantly enters the culture medium (Fig. 2f). These observations suggest that autoprocessing by fragment C may regulate the degree of N association with the cell surface and therefore its range of action.

The missense lesions within the class I group of mutations support the conclusion from our ectopic expression studies that N is responsible for both local and long-range Hh signalling. Given the retention of N near its cellular site of synthesis^{6,16,21}, this conclusion is less surprising for local Hh activities such as induction of *wg* and *dpp* expression in embryos and imaginal discs than it is for longer-range effects such as patterning of the dorsal cuticle. The *hh^{9k}* missense lesion in particular confirms a

and resemble those of denticle rows 2 or 3. k-o, Location and extent of four distinct cell types within a single segment of the dorsal cuticle of the first instar larva is summarized in the inset in each panel. In Hh (l) and Hh-N (n), secondary cell fates characterized by naked cuticle expand at the expense of all tertiary and some quaternary cell types. METHODS. To assay for the expansion of *wg* expression, embryos collected 4-6 h after egg laying (AEL) at 25 °C were subjected to two 15-min 37 °C heat shocks, separated by 90 min and allowed to recover for 40 min at 25 °C. The UAS N/hsGAL4 flies were given an additional hour at 25 °C to allow for the secondary induction of UAS N. *In situ* hybridizations were performed as described²⁵ using a *wg*-specific probe¹³. Embryos assayed for cuticle phenotypes were heat-shocked for 30 min at 37 °C at 5-8 h AEL, allowed to develop at 25 °C for 36 h and then processed and mounted as described²⁶.

long-range role for N, because this was the thermosensitive allele used to inactivate long-range dorsal cuticle patterning activity⁷. Models accounting for long-range Hh activities in *Drosophila* must now incorporate a signal relay mechanism involving induction of other signalling molecules or, alternatively, the existence of a steep gradient with active N present at levels not yet detectable by current techniques. Vertebrate Hh has been implicated in limb patterning^{12,13}, in local induction of floor plate by the notochord¹¹, and in long-range induction of sclerotome from presomitic mesoderm^{14,15}. Now that we have characterized the autoproteolytic cleavage site in vertebrate proteins, it should be possible to determine whether local and long-range Hh signalling functions in vertebrates are executed by N, as we have found in *Drosophila*. □

Received 27 December 1994; accepted 10 February 1995.

1. Nüsslein-Volhard, C. & Wieschaus, E. *Nature* **287**, 795-801 (1980).
2. Mohler, J. *Genetics* **120**, 1061-1072 (1988).
3. Ma, C., Zhou, Y., Beachy, P. A. & Moses, K. *Cell* **75**, 927-938 (1993).
4. Heberlein, U., Wolff, T. & Rubin, G. *Cell* **75**, 913-926 (1993).
5. Ingham, P. W. *Nature* **366**, 560-562 (1993).
6. Tabata, T. & Kornberg, T. B. *Cell* **76**, 89-102 (1994).
7. Heemskerk, J. & DiNardo, S. *Cell* **76**, 449-460 (1994).
8. Basler, K. & Struhl, G. *Nature* **368**, 208-214 (1994).
9. Echeiardi, Y. et al. *Cell* **75**, 1417-1430 (1993).
10. Krauss, S., Concordet, J. P. & Ingham, P. W. *Cell* **75**, 1431-1444 (1993).
11. Roelink, H. et al. *Cell* **76**, 761-775 (1994).
12. Riddle, R. D., Johnson, R. L., Laufer, E. & Tabin, C. *Cell* **75**, 1401-1416 (1993).
13. Chang, D. T. et al. *Development* **120**, 3339-3353 (1994).
14. Johnson, R. L., Laufer, E., Riddle, R. D. & Tabin, C. *Cell* **79**, 1165-1174 (1994).
15. Fan, C.-M. & Tessier-Lavigne, M. *Cell* **79**, 1175-1186 (1994).

16. Lee, J. J. et al. *Science* **266**, 1528-1537 (1994).
17. Jürgens, G., Wieschaus, E., Nüsslein-Volhard, C. & Kluding, H. *Roux's Arch. dev. Biol.* **193**, 283-295 (1984).
18. Vanderslice, P., Copeland, W. C. & Robertus, J. D. *J. biol. Chem.* **263**, 10583-10586 (1988).
19. Koonin, E. V. *Trends biochem. Sci.* (in the press).
20. Xu, M.-Q. et al. *EMBO J.* **13**, 5517-5522 (1994).
21. Taylor, A. M., Nakano, Y., Mohler, J. & Ingham, P. W. *Mech. Dev.* **42**, 89-96 (1993).
22. Rubin, G. M. & Spradling, A. C. *Science* **218**, 348-353 (1982).
23. Thummel, C. S., Boulet, A. M. & Lipshitz, H. D. *Gene* **74**, 445-456 (1988).
24. Brand, A. H. & Perrimon, N. *Development* **118**, 401-415 (1993).
25. Tautz, D. & Pfeiffe, C. *Chromosoma* **98**, 81-85 (1989).
26. Ashburner, M. *Drosophila: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, New York, 1989).

ACKNOWLEDGEMENTS. We thank W.-S. Liu for assistance with protein sequencing; J. Ptak and C. Davenport for oligonucleotide synthesis and automated DNA sequencing; and J. F. Maher and other colleagues for discussion. P.A.B. is an Assistant Investigator of the Howard Hughes Medical Institute.

Neural progenitor cell engraftment corrects lysosomal storage throughout the MPS VII mouse brain

Evan Y. Snyder*, Rosanne M. Taylor†‡
& John H. Wolfe†§

* Departments of Neurology and Pediatrics, Harvard Medical School, Children's Hospital, Boston, Massachusetts 02115, USA

† Laboratory of Pathology and Section of Medical Genetics, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

MANY metabolic diseases affecting the central nervous system are refractory to treatment because the blood-brain barrier restricts entry of therapeutic molecules. It may be possible to deliver therapeutic gene products directly to the brain by transplantation of neural progenitor cells, which can integrate into the murine central nervous system in a cytoarchitecturally appropriate manner¹⁻⁷. We tested this approach in mucopolysaccharidosis VII (Sly disease), a lysosomal storage disorder of humans, dogs and mice caused by an inherited deficiency of β -glucuronidase⁸⁻¹⁰. Lysosomal accumulation of glycosaminoglycans occurs in the brain and other tissues, causing a fatal progressive degenerative disorder, including mental retardation¹¹. Treatments are designed to provide a source of normal enzyme for uptake by diseased cells¹²⁻²⁰. We report here that by transplanting β -glucuronidase-expressing neural progenitors into the cerebral ventricles of newborn mice, donor cells engrafted throughout the neuraxis. At maturity, donor-derived cells were present as normal constituents of diverse brain regions. β -Glucuronidase activity was expressed along the entire neuraxis, resulting in widespread correction of lysosomal storage in neurons and glia in affected mice.

The multipotent C17.2 immortalized neural progenitor cell line^{1-4,21,22}, which was derived from the external germinal layer of neonatal mouse cerebellum, was implanted into the lateral ventricles of newborn mice. The C17.2 cells can differentiate into neurons or glia, in a manner appropriate to the site of engraftment, when transplanted into various structures of the brain and at various stages of development from fetus to adult¹⁻⁴. The engrafted cells reside on the brain side of the blood-brain barrier and develop into integral members of the central nervous system (CNS) cytoarchitecture, for example, donor-derived neurons receive synapses; the blood-brain barrier remains intact where donor-derived astroglia put foot processes onto cerebral vasculature; and donor-derived oligodendrocytes express myelin basic protein and myelinate¹⁻⁴. Injecting the progenitor cells into the lateral ventricles presumably allowed them to gain access to most of the subventricular germinal zone²³, including that of the IIIrd and IVth ventricles, as well as to networks of cerebral vasculature along the surface of which they may also have migrated. Recipients survived the procedure without cerebrospinal fluid obstruction or other morbidity, and no tumours were ever observed in these mice nor in hundreds of others, some of which were observed for up to 2 years¹⁻⁴.

Donor cells migrated into the parenchyma from the ventricles within 24 h of transplantation (not shown). Beginning at 3 weeks after transplantation, recipient brains were analysed for the presence of donor-derived cells by the presence of X-gal staining (5-bromo-4-chloro-indoyl precipitate) for β -galactosidase activity expressed from a *lacZ* marker gene¹. By 3 weeks of age, the pathological phenotype is fully expressed in the MPS VII mouse brain^{10,24} and CNS organogenesis is complete. X-gal-positive

cells were found throughout the recipient brains (Fig. 1). All but one of the MPS VII animals (16/17) had evidence of successful engraftment, though individual recipients varied in the amount of engraftment (Table 1) as seen previously¹. Donor cells were present at all time points post-transplantation (Table 1), indicating that engraftment was stable, and probably was permanent because donor cells have been identified in normal recipients for at least 22 months¹.

The donor-derived cells expressing *lacZ* assumed normal neural morphologies in all regions examined, suggesting that the transplanted cells differentiated into integral members of the CNS cytoarchitecture. The cells intermingled non-disruptively with host cells. The X-gal-positive cells were present in non-symmetrical patterns (Fig. 1), indicating that variations in migration, expansion and differentiation of the progenitors had occurred. There were no apparent differences in engraftment between normal or MPS VII recipients. Furthermore, long-term engrafted animals (6-12 months) exhibited no behavioural abnormalities or other indications of neurological dysfunction. Thus the structures of the brain that received contributions from donor cells appeared to have developed normally.

The treated MPS VII brains were examined for localization of β -glucuronidase (GUSB) activity using a histochemical enzymatic reaction²⁵ in sections adjacent to those containing X-gal-positive cells. The MPS VII mouse has a deletion mutation²⁶ and is devoid of GUSB activity^{10,13-20} (for example, see Fig. 3A, a), thus any detectable GUSB activity was from donor cells. The most intense GUSB staining colocalized with X-gal-positive regions (Fig. 1). Sections throughout the neuraxis of the brains were positive, and GUSB activity was absent in areas that do not receive contributions from the subventricular zone, for example, the retina (not shown).

The quantitative distribution of GUSB activity in the recipient brains was determined from representative samples from regions of each brain divided coronally as shown in Fig. 2. All of the engrafted animals had evidence of elevated GUSB activity in the brain, including normals and carriers (Table 1). There was no evidence of any adverse effect of GUSB overexpression on these animals, consistent with previous findings in transgenic mice²⁷ and transduced cells²⁸. The donor-derived GUSB activity was present in diverse regions throughout the brains of all recipients (Table 1). There was no apparent predilection for engraftment in any specific region because the highest levels in individual brains were found in any of the five regions, and the average activities by region for the treated mice were similar (Table 1). The variations in total GUSB activity in individual MPS VII recipients (Table 1) were probably due to variations in engraftment because the mice that had lower levels of GUSB expression had correspondingly fewer X-gal-positive cells (not shown).

Of the transplanted MPS VII mice, 94% had evidence of diffuse engraftment, with an average of at least 2% of normal activity (Table 1), which can be corrective¹⁴. Many regions expressed at least 1.4 units ($\text{nmol h}^{-1} \text{mg}^{-1}$) of GUSB activity, an amount that reverses pre-existing storage in the liver and spleen¹⁴. Some areas approached carrier levels (Table 1). The normal lysosomal enzyme α -galactosidase, which is nonspecifically elevated in some tissues in MPS VII and can return towards normal after correction of the GUSB deficiency^{13,14,16}, was closer to normal in the brains of most of the treated mutants (not shown). There was significant correlation between treatment at birth and the presence of GUSB activity in later life ($P < 0.001$), and between the presence of X-gal-positive and GUSB-positive cells ($P < 0.005$).

Lysosomal distention is present in the MPS VII mouse brain at birth and reaches its full extent by three weeks of age²⁴. Because engraftment of donor cells appears to be permanent (ref. 1 and Table 1), treated MPS VII mice were analysed at 3 weeks post-transplantation for differences in lysosomal storage compared

‡ Present address: Department of Animal Science, University of Sydney, Sydney, Australia.

§ To whom correspondence should be addressed.

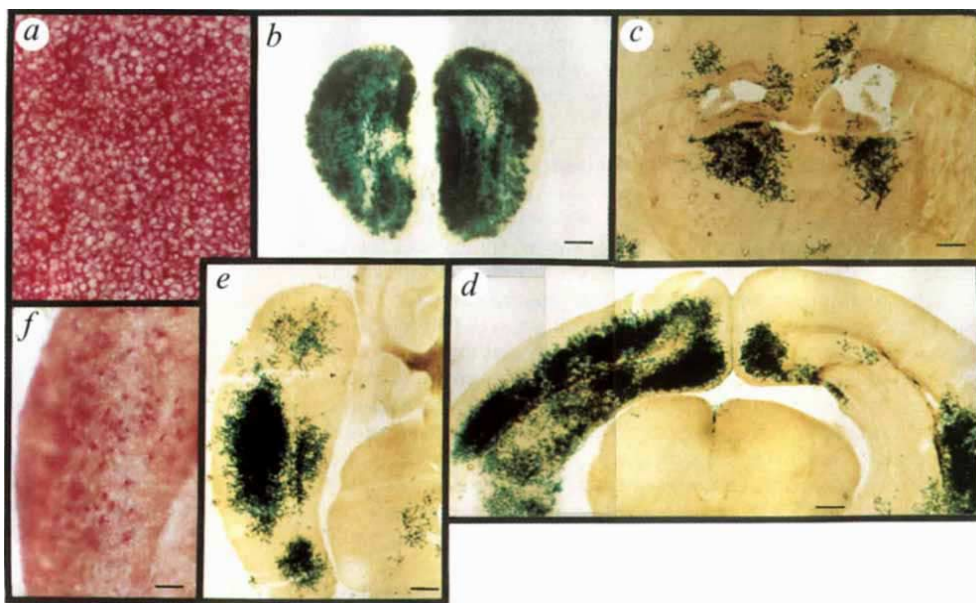
to untreated controls. The cytoplasmic vacuolation characteristic of lysosomal storage was substantially decreased or absent from both neurons and glia. Dramatic improvements were seen in all engrafted areas of the transplanted mutant brains examined, including the cerebral cortex (Fig. 3A). To confirm these results, animals nearing the ends of their lives (>5 months¹⁰) were also examined and compared to age-matched untreated controls, with similar results (Fig. 3B). As is routinely seen in successfully treated visceral organs^{13,14,16,20}, some storage could

be found in the treated brains. There was a significant correlation between treatment and changes in storage ($P < 0.001$).

Rescued glia that were X-gal-negative were presumably of host origin. The absence of storage seen in some glial cells could, however, be attributed to replacement of host cells by normal donor-derived cells rather than correction of mutant cells. This possibility derives from earlier experiments demonstrating that donor progenitors respond to the prevalent spatial and temporal developmental cues in the area of engraftment¹⁻⁴. Progenitors

FIG. 1 Brain of a mature MPS VII mouse after receiving a neonatal intraventricular transplant of C17.2 neural progenitor cells expressing β -glucuronidase (GUSB). **a**, GUSB staining (red reaction product²⁵) of C17.2 cells in culture before transplantation. **b–e**, Identification *in vivo* of donor-derived cells by X-gal histochemical reaction (blue precipitate) for expression of the *lacZ* marker gene. The blue cells can be seen to have engrafted throughout the recipient mutant brain (MPS #84, Table 1). Representative regions are shown, proceeding clockwise from rostral-to-caudal: **b**, olfactory bulbs; **c**, telencephalon at the level of the caudal (including striatum, corpus callosum and fronto-parietal cortex); **d**, telencephalon at the level of the rostral aspect of the hippocampus (including parieto-occipital cortex and hippocampus); and **e**, posterior telencephalon (including occipital cortex) and midbrain. **f**, Expression of GUSB (red cells) in a sister section to **e**, showing colocalization of X-gal-positive (blue) and GUSB-positive (red) donor cells. Untreated MPS VII mice show no GUSB activity histochemically^{13–16,19} (see Fig. 3). Scale bars, 400 μ m (**b–e**) and 320 μ m (**f**).

METHODS. C17.2 cells were maintained and prepared for transplantation as previously described¹. The heads of cryoanesthetized newborn mice were transilluminated and 2 μ l containing $2-8 \times 10^4$ cells plus 0.05% w/v trypan blue in PBS were gently expelled into each lateral ventricle through a finely drawn glass micropipette. The distribution of trypan blue confirmed filling of the ventricular system from as far rostral as the olfactory bulbs to as far caudal as the IVth ventricle and the rostral central spinal canal. The pups were gently warmed and returned to maternal care when active. All pups in litters from carrier (MPS/+ $\times$ MPS/+) matings were transplanted, and their genotypes



determined later¹⁰ as mortality from the procedure was low. Recipients were killed at various times post-transplantation and the brains sliced coronally at 20–100- μ m serial intervals. Sections were processed for *lacZ* expression by X-gal histochemistry¹, for GUSB activity *in situ*^{14,15}, and for plastic embedding¹ to assess lysosomal storage following toluidine blue staining¹⁴ (see Fig. 3). Sister sections, not used for X-gal or GUSB staining, were assayed for quantitative GUSB enzymatic activity^{10,14} (Table 1). The C17.2 cells produced ~ 800 units ($\text{nmol h}^{-1} \text{mg}^{-1}$) of mouse GUSB activity *in vitro*. The cells contained retroviral vectors with *lacZ*¹, *myc*¹ and human GUSB²⁸ cDNAs, were free of helper virus before transplantation^{1,22}, and these sequences were detected by PCR as additional markers of donor cell engraftment for at least a year post-transplant in normal recipients (not shown).

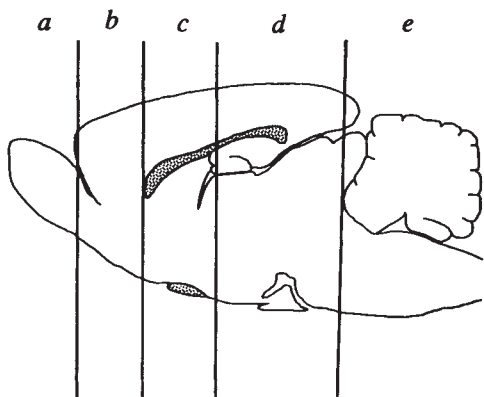


FIG. 2 Scheme for assessing distribution of GUSB enzymatic activity in transplanted brains. Serial sister coronal sections that were not used for X-gal or GUSB histochemical staining were collected from throughout the brains of transplant recipients and quantitatively assayed for GUSB activity^{10,14}. They were pooled to reflect the activity present within the regions demarcated in this schematic and used in Table 1. The regions were defined by anatomical landmarks in the anterior-to-posterior plane to permit comparison among animals: **a**, olfactory bulbs; **b**, caudal edge of the olfactory bulbs to the beginning of the striatum; **c**, striatum to the rostral edge of the hippocampus; **d**, hippocampus to the posterior colliculus; **e**, posterior colliculus of the midbrain (just rostral to cerebellum) to the beginning of the spinal cord. The cerebellum was dissected free and was not included in the assay, thus region **E** largely reflected the activity in the brainstem. Examples of the distribution and activity of engrafted cells within representative coronal sections from MPS VII mice can be seen from GUSB histochemistry in Figs 1 and 3.

engrafted within the cerebral cortex at birth give rise almost exclusively to glia²⁻⁴ because neurogenesis has ceased in this region by that time but gliogenesis is still occurring. Thus it is most likely that corrected neurons within the cortex of the treated MPS VII mice were predominantly of host origin. Accordingly, normal X-gal-negative neurons were found in areas of transplanted MPS VII mice which show substantial amounts of storage in untreated mutants (Fig. 3). The corrected host cells probably incorporated normal GUSB from the donor-derived cells by mannose-6-phosphate receptor-mediated endocytosis (cross-correction), which is the basis for treatment of MPS VII and most other lysosomal storage diseases^{11,29}.

The neonatal transplantation of GUSB-producing neural progenitors corrected the disease in many areas of the MPS VII mouse brain, as neonatal bone marrow transplantation did for

the visceral and skeletal disease¹⁶. Neural progenitor cell transplantation did not require pretransplantation conditioning of recipients (such as irradiation), which is deleterious to the CNS in newborn mice and in children receiving bone marrow transplants^{16,30}. Therapy instituted early in life might arrest disease progression in many inherited metabolic diseases. This may be especially important for treating diseases causing mental retardation because it is unclear whether permanent alterations in brain function occur during early life that may not be reversible even if normal metabolism is restored later.

The approach we have described provides a model for using neural progenitors to transfer other exogenous genes or factors into the CNS. Although many methods are currently being studied to deliver foreign gene products to the brain³¹⁻³⁴, exploiting the inherent biological properties of neural progenitors to

FIG. 3 A, Decreased lysosomal storage throughout the brains of MPS VII recipients of C17.2 transplants.

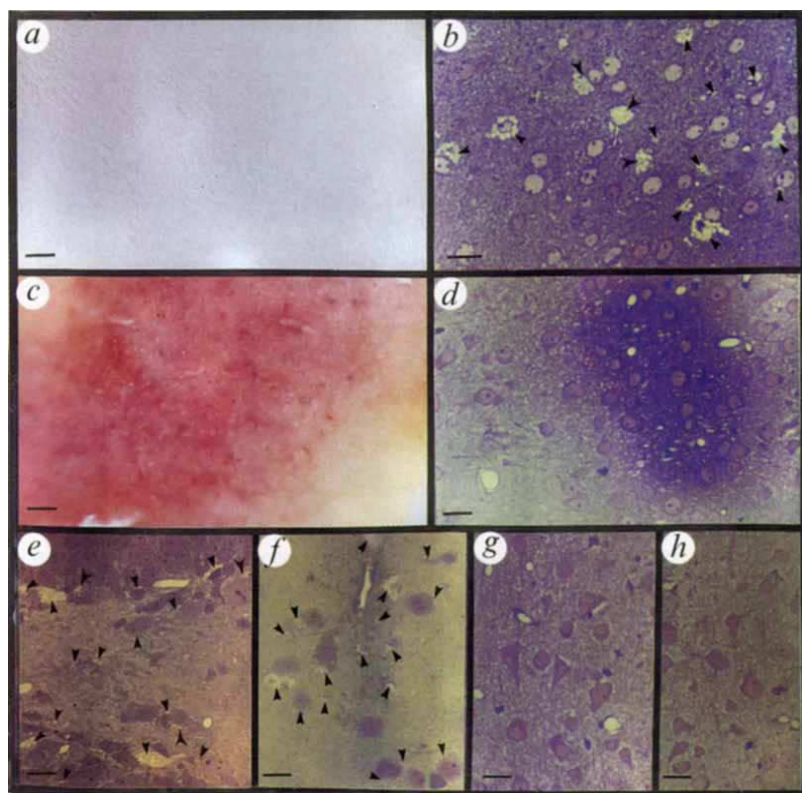
Lysosomal storage was examined in regions demarcated in Fig. 2 by toluidine blue staining of X-gal-processed, plastic-embedded sections, and found to be absent or diminished compared to analogous regions of untransplanted, age-matched MPS VII controls. Representative regions are shown, but the other regions of treated brains showed similar decreases in lysosomal storage. *a-d*, Donor cell GUSB activity and reduction of storage in the hippocampus (from region *d* in Fig. 2) at 3 weeks post-transplant. *a*, Hippocampus of a control untransplanted 3-week-old MPS VII mouse was completely negative for GUSB histochemical reaction product and *b* showed abundant white cytoplasmic vacuoles representing distended lysosomes (arrowheads). *c*, The same region from a 3-week-old MPS VII mutant (#238, Table 1) transplanted neonatally with C17.2 cells. The dense red stain represents extensively distributed normal GUSB expression by engrafted cells. *d*, A toluidine-blue-stained, plastic-embedded section of the hippocampus from mouse #238, adjacent to the GUSB-positive section shown in *c* and analogous to the untreated control section in *b*, showing the absence of white cytoplasmic vacuoles in neurons and glia where GUSB is secreted. The large white holes are blood vessels. The section had previously been processed with X-gal histochemistry.

e-h, Decreased lysosomal storage in the neocortex (from region *c*, Fig. 2) of the same MPS VII mouse #238. *e, f*, Toluidine-blue-stained cortex and subcortical area of an untransplanted 3 week old MPS VII mouse showing abundant white cytoplasmic vacuoles representing distended lysosomes (arrowheads). *g, h*, Toluidine-blue-stained, X-gal-processed sections from treated mouse #238 from an analogous region as shown in the control, untreated mouse (*e, f*). Little if any lysosomal storage is seen in either neurons or glia, which are probably cross corrected cells of host origin because they are X-gal-negative and because many of the corrected cells are neurons, a cell type into which progenitors in a postnatal neocortex would not differentiate. Scale bars, 320 μ m (*a, c*); 31 μ m (*b, d, g, h*); 25 μ m (*e, f*).

B, Decreased lysosomal storage in a treated brain at 8 months of age. *a*, Extensive vacuolation representing distended lysosomes (arrowheads) in both neurons and glia in the cortex of an 8-month-old, untransplanted control MPS VII mouse (from region *c* in Fig. 2). *b, c*, Decrease in lysosomal storage in neurons and glia in the neocortex of MPS VII mouse #192 in an X-gal-processed section from a region analogous to the untreated control section *a*. The other regions of this animal's brain showed a similar decrease in lysosomal storage compared to untreated, age-matched mutants in regions where GUSB was expressed. Toluidine blue staining was done as in Fig. 3. Scale bars, 21 μ m (*a, c*) and 31 μ m (*b*).

METHODS. GUSB staining (*a, c*) was as previously described^{14,15}. Toluidine blue staining (*a, d, e-h*) was done on 0.5- μ m-thick sections cut from brain slices processed with X-gal histochemistry, embedded in epon-alardite, and photographed under brightfield as previously described^{11,14}.

A



B

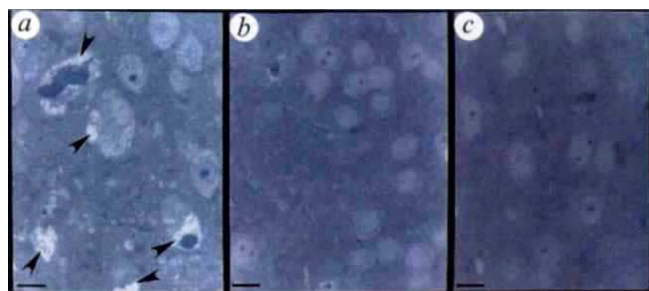


TABLE 1 Expression of β -glucuronidase in brains of mice transplanted at birth with C17.2 neural progenitor cells

Treatment/genotype	Age at analysis*	GUSB staining†	Whole brain average	GUSB enzyme activity (in nmol h ⁻¹ mg ⁻¹ (+s.e.) (% of average normal level))				
				Region				
				A	B	C	D	E
Untransplanted wild types (n=3)	3 w	+	27.3 (2.7) (100%)‡	28.4 (2.3) (100%)	27.7 (3.9) (100%)	26.8 (3.1) (100%)	27.6 (3.7) (100%)	25.8 (0.5) (100%)
Untransplanted carriers (n=6)	3 w–10 m	+	16.8 (1.1) (62%)	16.1 (1.4) (57%)	13.4 (0.7) (48%)	14.4 (0.6) (54%)	15.6 (1.0) (57%)	24.6 (1.8) (95%)
Untransplanted MPS VII (n=10)	3 w–8 m	—	0.07 (0.02) (0%)§	0.01 (0.01) (0%)	0.05 (0.03) (0.2%)	0.07 (0.03) (0.3%)	0.08 (0.03) (0.3%)	0.03 (0.01) (0%)
Transplanted wild types (n=7)	3 w–10 m	+	46.9 (3.7) (172%)	58.8 (6.7) (208%)	40.5 (3.7) (146%)	52.8 (6.6) (197%)	35.2 (1.9) (128%)	47.2 (2.3) (183%)
Transplanted carriers (n=5)	3 w–10 m	+	23.7 (2.7) (87%)	19.4 (0.3) (69%)	21.4 (1.5) (78%)	22.4 (3.1) (84%)	27.1 (4.8) (98%)	27.9 (1.5) (108%)
Transplanted MPS VII								
MPS 72	1 d	+	12.4 (7.1) (45%)	15.3 (55%)	15.3 (55%)	20.1 (75%) 	1.7 (6%)	1.7 (6%)
MPS 84	3 w	+	5.5 (4.7) (20%)	10.9 (38%)	0.8 (3%)	2.6 (10%)	11.9 (43%)	1.4 (5%)
MPS 201	3 w	—#	0.6 (2%)**	NA	NA	NA	NA	NA
MPS 202	3 w	+	0.6 (2%)**	NA	NA	NA	NA	NA
MPS 208	3 w	+	1.2 (1.0) (4%)	2.6 (9%)	2.3 (8%)	0.2 (1%)	0.1 (0.4%)	0.6 (2%)
MPS 207	3 w	no engraft*	0.0 (0.0) (0%)	0.0 (0%)	0.0 (0%)	0.0 (0%)	0.0 (0%)	0.0 (0%)
MPS 238	3 w	+	4.1 (1.3) (14%)	3.9 (14%)	2.7 (10%)	6.7 (25%)	2.9 (11%)	2.9 (11%)
MPS 239	3 w	+	7.2 (3.2) (26%)	9.9 (35%)	11.0 (40%)	1.9 (7%)	6.1 (22%)	6.1 (22%)
MPS 240	3 w	+	1.0 (0.5) (4%)	0.3 (1%)	2.1 (8%)	0.9 (3%)	0.7 (3%)	0.8 (3%)
MPS 223	1 m	+	0.4 (0.2) (2%)	0.6 (2%)	0.3 (1%)	0.2 (1%)	0.7 (3%)	0.2 (1%)
MPS 222	2 m	+	0.5 (0.0) (2%)	0.6 (2%)	0.6 (2%)	0.5 (2%)	0.5 (2%)	0.5 (2%)
MPS 221	2 m	+	0.4 (0.1) (2%)	0.7 (3%)	0.4 (1%)	0.3 (1%)	0.4 (2%)	0.4 (2%)
MPS 220	2 m	+	0.7 (0.4) (3%)	1.8 (6%)	0.6 (2%)	0.7 (3%)	0.3 (1%)	0.3 (1%)
MPS 236	3 m	ND	0.6 (0.1) (2%)	0.6 (2%)	0.7 (3%)	0.5 (2%)	0.6 (2%)	0.5 (2%)
MPS 244	4 m	+	0.7 (0.0) (3%)	0.7 (3%)	0.7 (3%)	0.6 (2%)	0.7 (3%)	0.6 (2%)
MPS 256	6 m	+	9.8 (12.9) (38%)	2.0 (7%)	0.0 (0%)	5.1 (19%)	0.0 (0%)	42.1 (163%)
MPS 192	8 m	+	2.4 (0.7) (9%)	3.4 (12%)	1.5 (5%)	1.9 (7%)	2.8 (10%)	2.8 (10%)
Average of all engrafted MPS mice			3.04(2.3) (11%)	4.4 (15.5%)	3.6 (12.9%)	3.9 (14.6%)	2.6 (9.5%)	6.6 (25.3%)

MPS, Mucopolysaccharidosis type VII; wk, week; mo, month; NA, not applicable; ND, not determined; GUSB, β -glucuronidase; s.e., standard error.

* Age at analysis is the same as time post-transplant.

† All wild-type and carrier (heterozygote) mice stain positively for GUSB activity throughout the brain and MPS VII mice are always completely negative (see Fig. 3). GUSB-positive cells colocalized with engrafted XGal-positive donor-derived cells in transplanted mutants.

‡ The mean GUSB activity levels for entire brain and for each region of untransplanted wild-type mice represent the 'normal GUSB activity' for this strain. This value serves as the denominator for calculating the 'per cent of normal activity' for whole brain and for specific regions, respectively, in the other categories¹⁰. Similarly, the enzyme activity levels values for untreated carrier (heterozygote), and untreated mutant (homozygote) MPS VII mice, reflect their baseline values respectively.

§ These values represent the background reading of the assay and not residual activity. As previously stated and established in the literature¹⁰, homozygotes have no GUSB activity.

|| In MPS 72 regions A and B, and D and E, were pooled and assayed together because of the small size of the brain. In MPS 238, 239 and 192, sections from regions D and E were pooled.

¶ The region with the highest GUSB activity for each engrafted MPS animal is indicated in bold. Note that this region varies randomly from animal to animal.

In this animal, low numbers of β -gal-positive cells were detected but nearby sections were negative for GUSB-positive staining, which was probably a sampling artefact because there was measurable GUSB activity in other sections assayed quantitatively.

* This is the only transplanted mutant in this study that failed to engraft, yielding a 94% transplant efficiency. All engrafted animals had significant, measureable GUSB activity. This negative animal was not included in calculations of average enzyme activities for engrafted animals.

** These brains were sliced axially and not coronally. Representative sections were pooled for a single determination of total brain GUSB activity. Therefore, distribution of GUSB could not be determined as in Fig. 2, but histological evidence of engraftment was evident throughout the brain in these 2 animals.

become integral members of normal structures throughout the host brain may allow foreign gene products to be delivered in a sustained, direct and perhaps regulated fashion, without disturbing other neurobiological processes. □

Received 23 August 1994; accepted 16 January 1995.

- Snyder, E. Y. *et al.* *Cell* **68**, 33–51 (1992).
- Snyder, E. Y. *Semin. Perinatol.* **16**, 106–121 (1992).
- Snyder, E. Y., Yandava, B. D., Pan, Z.-H., Yoon, C. & Macklis, J. D. *Soc. Neurosci. Abstr.* **19**, 613 (1993).
- Macklis, J. D., Yoon, C. H. & Snyder, E. Y. *Expl. Neurol.* **129**, 9 (1994).
- Renfranz, P. J., Cunningham, M. G. & McKay, R. D. G. *Cell* **66**, 713–719 (1991).
- Groves, A. K. *et al.* *Nature* **362**, 453–455 (1993).
- Onnifer, S. M., Whittemore, S. R. & Holets, V. R. *Expl. Neurol.* **122**, 130–142 (1993).
- Sly, W. S., Quinton, B. A., McAlister, W. H. & Rimoin, D. L. *J. Pediatr.* **82**, 249–257 (1973).
- Haskins, M. E., Desnick, R. J., DiFerrante, N., Jezyk, P. F. & Patterson, D. F. *Pediatr. Res.* **18**, 980–984 (1984).
- Birkenmeier, E. H. *et al.* *J. clin. Invest.* **83**, 1258–1266 (1989).
- Neufeld, E. F. & Muenzer, J. in *The Metabolic Basis of Inherited Disease* 7th edn (eds Scriver, C. R., Beaudet, A. L., Sly, W. & Valle, D.) 2465–2494 (McGraw Hill, New York, 1995).
- Haskins, M. E. *et al.* in *Treatment of Genetic Diseases* (ed. Desnick, R. J.) 183–202 (Churchill Livingstone, New York, 1991).
- Birkenmeier, E. H. *et al.* *Blood* **78**, 3081–3092 (1991).
- Wolfe, J. H. *et al.* *Nature* **360**, 749–753 (1992).
- Wolfe, J. H., Deshmane, S. L. & Fraser, N. W. *Nature Genet.* **1**, 379–384 (1992).
- Sands, M. S. *et al.* *Lab. Invest.* **68**, 676–686 (1993).

- Moullier, P., Bohl, D., Heard, J.-M. & Danos, O. *Nature Genet.* **4**, 154–159 (1993).
- Marechal, V., Naffakh, N., Danos, O. & Heard, J.-M. *Blood* **82**, 1358–1365 (1993).
- Vogler, C. *et al.* *Pediatr. Res.* **34**, 837–840 (1993).
- Sands, M. S. *et al.* *J. clin. Invest.* **93**, 2324–2331 (1994).
- Ryder, E. F., Snyder, E. Y. & Cepko, C. L. *J. Neurobiol.* **21**, 356–375 (1990).
- Kitchens, D. L., Snyder, E. Y. & Gottlieb, D. I. *J. Neurobiol.* **25**, 797–807 (1994).
- Altman, J. *J. comp. Neurol.* **137**, 433–458 (1969).
- Vogler, C. *et al.* *Am. J. Path.* **136**, 207–217 (1990).
- Wolfe, J. H. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **87**, 2877–2881 (1990).
- Sands, M. S. & Birkenmeier, E. H. *Proc. natn. Acad. Sci. U.S.A.* **90**, 6567–6571 (1993).
- Kyle, J. W. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **87**, 3914–3918 (1990).
- Wolfe, J. H. *et al.* *Gene Ther.* **2**, 70–78 (1995).
- Taylor, R. M. & Wolfe, J. H. *Expl. Cell Res.* **214**, 606–613 (1994).
- Christie, D. *et al.* *Arch. Dis. Child* **70**, 275–280 (1994).
- Suhr, S. & Gage, F. H. *Arch. Neurol.* **50**, 1252–1268 (1993).
- Breakfield, X. O. *Nature Genet.* **3**, 187–189 (1993).
- Fisher, L. J. & Gage, F. H. *FASEB J.* **8**, 489–496 (1994).
- Anton, R. *et al.* *Expl. Neurol.* **127**, 207–218 (1994).

ACKNOWLEDGEMENTS. We thank B. Yandava, A. Polesky, M. Parente and S. Cassel for technical assistance, J. Comander for assistance with statistical analysis and M. Sands and J. Flax for advice. This work was supported by grants from the National Institute of Neurological Disorders and Stroke, the Charles H. Hood Foundation, the William Randolph Hearst Foundation, the American Paralysis Association and the Paralyzed Veterans of America (EYS); and from the National Institute of Diabetes and Digestive and Kidney Diseases, the Lucille P. Markey Charitable Trust and the Mrs Cheever Porter Foundation (JHW). The Mental Retardation Research Centers at Harvard University and the University of Pennsylvania and the Gene Therapy Core Center at the University of Pennsylvania provided technical support. E.Y.S. was the recipient of a Clinical Investigator Development Award from the National Institute of Neurological Disorders and Stroke, and R.M.T. was supported by a Kleberg Fellowship in Veterinary Medical Genetics and a Sandoz Fellowship from the Transplantation Society of Australia and New Zealand.

DNA-bend modulation in a repressor-to-activator switching mechanism

Aseem Z. Ansari, James E. Bradner & Thomas V. O'Halloran*

Department of Chemistry and Department of Biochemistry, Molecular Biology and Cell Biology, 2145 Sheridan Road, Northwestern University, Evanston, Illinois 60208-3113, USA

* To whom correspondence should be addressed

RECENT discoveries of activator proteins that distort DNA but bear no obvious activation domains have focused attention on the role of DNA structure in transcriptional regulation¹. Here we describe how the transcription factor MerR can mediate repression as well as activation through stereospecific modulation of DNA structure. The repressor form of MerR binds between the -10 and -35 promoter elements of the bacterial mercury-detoxification genes, P_T , allowing RNA polymerase to form an inactive complex with P_T and MerR at this stress-inducible promoter^{2,3}. Upon mercuric ion binding, Hg-MerR converts this polymerase complex into the transcriptionally active or 'open' form²⁻⁴. We show here that MerR bends DNA towards itself in a manner similar to the bacterial catabolite-activator protein CAP, namely at two loci demarked by DNase I sensitivity, and that the activator conformation, Hg-MerR, relaxes these bends. This activator-induced unbending, when coupled with the previously described untwisting of the operator⁵, remodels the promoter and makes it a better template for the poised polymerase.

As inferred from the proximity of RNA polymerase and MerR footprints at P_T (refs 2-4), protein-protein contacts are possible but have yet to be established for repression or activation in this system⁶. We use phase-bending and DNase I assays here to test another mechanism in which transmission of the activation signal from regulatory protein to polymerase is mediated through a change in DNA structure. In one model, Hg-MerR underwinds the 19-base-pair (bp) spacer DNA and thereby optimizes the dihedral angle between suboptimally phased -10 and -35 elements of the P_T promoter (Fig. 1). Although a variety of experiments support an Hg-MerR-induced distortion of the spacer²⁻⁹, experiments described here establish that Hg-MerR

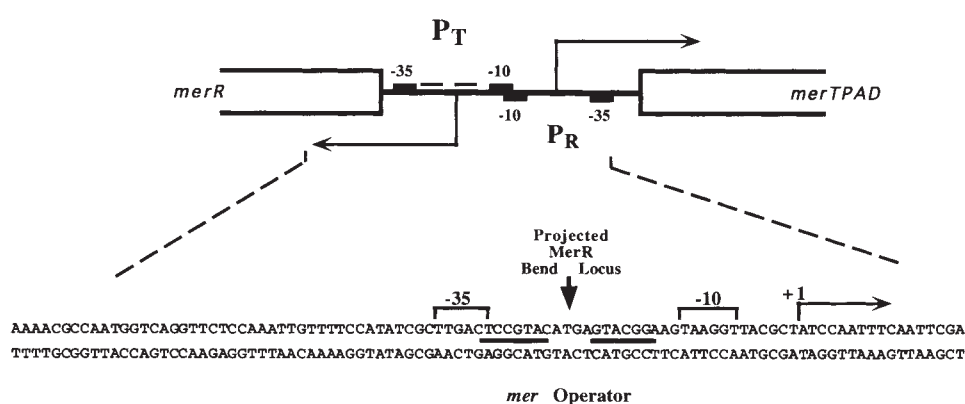
can reorient the trajectory of the -10 and -35 elements in the repressor-to-activator conversion.

In phasing assays, a known bend sequence is positioned at specific distances from an unknown bend and the phase-dependent variation in electrophoretic mobility reveals the relative direction of binding^{10,12}. This assay capitalizes on the observation that static bends in the same plane and direction reduce the distance between the ends of a DNA molecule, thus reducing its electrophoretic mobility¹². In this study, a 33-bp *mer* operator fragment is placed at incremental distances from a well characterized set of six phased A-tracts¹³, yielding a series of phasing constructs (Fig. 2a). The subtle pattern in the mobilities of free DNA fragments is sinusoidal and the amplitude of this pattern increases upon MerR complex formation (Fig. 2b), demonstrating that MerR induces a static bend in the operator. The helical distance between the centre of the operator and the projected locus of the reference bend is ~ 7.5 helical turns for the slowest-migrating complex (Fig. 2c). This half-integral value indicates that the reference and MerR-induced bend centres are on opposite faces of the DNA helix. As the bend direction at the projected centre of the reference bend is towards the major groove¹³, the bend direction at the projected centre of the MerR-induced bend is towards the minor groove. Footprinting data indicate that the minor groove faces MerR at the middle of the binding site⁴, leading to the conclusion that the protein bends the operator towards, not away from, itself.

Whereas Hg(II) alone does not significantly affect the migration of free DNA under these conditions (data not shown), the amplitude is completely lost upon formation of the Hg-MerR/DNA complex (Fig. 2). The Hg(II)-responsive decrease in the amplitude of the sine wave could result from a loss of the static bending or from changes in anisotropic flexibility at the MerR-induced bend sites^{10,14}. The DNase I cleavage profile provides insight into protein-induced distortions in DNA structure and localizes these structural changes to specific base steps. Hence we use it to distinguish between these two possibilities.

The MerR-DNA complex has a pair of DNase I cleavage sites within the protected region which are significantly diminished when Hg(II) is bound to MerR (Fig. 3). DNase I cleavage is known to be a function of backbone accessibility¹⁵, site flexibility¹⁶ and directed bending¹⁷, and we show here that DNase I-sensitive sites within protected regions of this and other nucleoprotein complexes are characteristic of a class of kinks in which DNA is bent towards the major groove. First, crystallographic studies of oligonucleotide-DNase I complexes reveal

FIG. 1 Two divergent overlapping promoters of the Tn501 *mer* operon⁴ are shown: P_T controls the expression of the detoxification genes and P_R is the promoter for the *merR* gene. MerR represses each promoter by a different mechanism, apparently blocking RNA polymerase (RNAP) from binding at P_R but allowing it to bind to P_T in an inactive ternary complex. Upon binding of Hg(II), MerR remains bound to the same site and still represses P_R but now is a potent activator of P_T . In the absence of *trans*-factors, P_T is a poor substrate for RNAP because its 19-bp spacer region deviates from the consensus length of 17-bp. Deletion to a 17-bp spacer yields an efficient promoter that is independent of, and is in fact weakly repressed by, Hg-MerR⁷. The P_T promoter sequence used in footprinting assays is shown with the -10 and -35 elements bracketed, the operator half-sites underlined, and $+1$ mark the transcription initiation site. Circular permutation experiments with the *mer* operator localize the $25^\circ \pm 10^\circ$ MerR-induced bend to the centre of the binding site (bold arrowhead) and little difference is observed upon formation of the Hg-MerR complex⁵; however, this assay does not discriminate between various DNA distortions as rigorously as



bend-phasing methods^{10,14}. Permutation and phasing analysis with the entire P_T promoter suggests the presence of weak intrinsic bends (A.Z.A. and T.V.O., unpublished results). Footprinting experiments on MerR and RNAP complexes reveal that Hg-MerR²⁻⁴ and allosteric control MerR^{AC} mutants that activate in the absence of Hg(II)^{6,8} each distort the sub-optimal spacer of P_T , whereas chemical nuclease experiments and DNA-linking-number assays are consistent with Hg-MerR-mediated underwinding of the spacer^{2,5}.

that the enzyme prefers bent substrates: it widens the minor groove by $\sim 15^\circ$, concomitantly bending DNA away from itself towards the major groove by $\sim 20^\circ$ (ref. 15). The scissile bond is one phosphodiester 3' to the site of the enzyme-induced kink¹⁵. Second, DNase I-hypersensitive sites within the CAP footprint on a symmetrized CC site¹⁸ bracket the $\sim 45^\circ$ CAP-induced bends¹⁹ that widen the minor groove (Fig. 3d). As shown in Fig. 3c, enhanced DNase I cleavage is observed at the phosphodiester immediately 3' to the CAP-induced kink¹⁸. A correlation between protein-induced bending and DNase I cleavage is also apparent in recently determined structures, including the Arc repressor-operator complex²⁰.

Comparison of DNase I cleavage profiles of crystallographically characterized complexes with that of MerR reveals a striking similarity with the CAP-DNA complex (Fig. 3c). Based on our data, and building on previous results^{2,4,5}, we propose that the repressor form of MerR bends DNA in a manner analogous to CAP, namely at two sites immediately 5' to the DNase I-sensitive sites, and that the activator form, Hg-MerR, unflexes these sites as it underwinds the centre of the operator. The Hg(II)-induced decrease in amplitude seen in the phasing experiments (Fig. 2) can originate from an increase in site flexibility or from a loss of static DNA bending. The decreased rate of DNase I cleavage in the Hg-MerR complex suggests that these sites are not flexible. We conclude that Hg-MerR decreases the bending of DNA relative to MerR. The fact that hydroxyl radical protection patterns remain unchanged with or without Hg(II) (Fig. 3b) indicates that a gross rearrangement of protein-

DNA contacts is not responsible for decreased DNase I cleavage at this site. The similarity of the hydroxyl radical cleavage patterns also demonstrates that the minor groove is not strongly compressed at these sites in either the repressed or activated states²¹. The mobility pattern observed in circular permutation studies of Hg-MerR⁵ may arise from other distortions. Bend-phasing methods are generally more rigorous probes of DNA bending than are circular permutation methods^{10,14}.

In Fig. 4a, we propose a model in which MerR-induced kinks occur at CpG base steps and are relaxed or lost upon conversion to the activator form. Studies of the steric properties of CpG (ref. 22), as well as systematic mutational analysis of CAP-DNA flanking sites, rank the CpG step as appropriate for a bend towards the major groove¹². Projection of these two symmetrically disposed bends toward the major groove is consistent with an average bend towards the minor groove at the centre of the binding site as described above.

Given the presence of obvious bends in the MerR/DNA complex that are diminished or lost upon Hg(II) binding, we examined the energetics of distortion as reflected in changes in the population linking number (ΔLk)⁵. Topological assays reveal that both the activator and repressor forms of MerR decrease the linking number:

$$\text{MerR} + \text{operator} \rightleftharpoons \text{MerR-operator} \quad \Delta Lk_1 \approx -0.05$$

$$\text{Hg-MerR} + \text{operator} \rightleftharpoons \text{Hg-MerR-operator} \quad \Delta Lk_2 \approx -0.15$$

$$\text{Hg(II)} + \text{MerR-operator} \rightleftharpoons \text{Hg-MerR-operator} \quad \Delta Lk_3 \approx -0.10$$

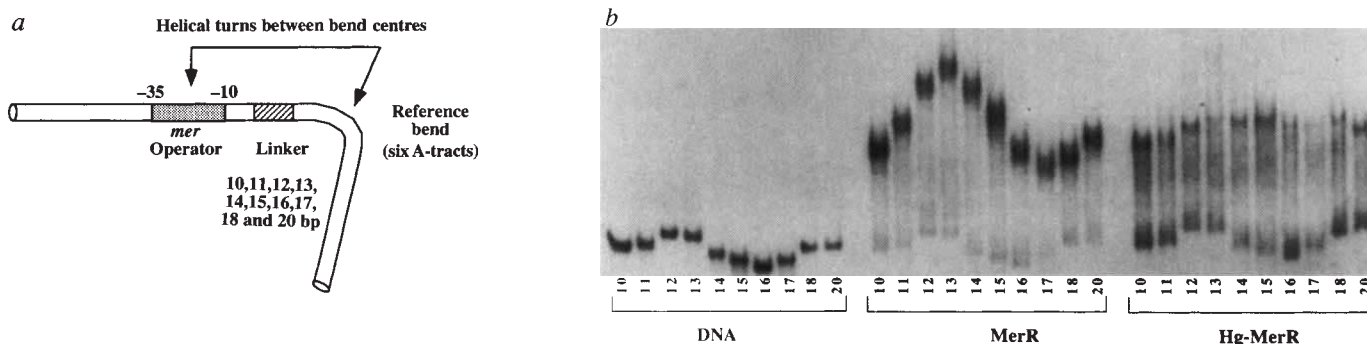
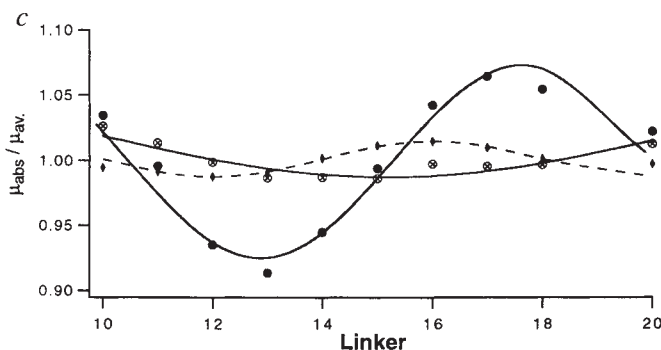


FIG. 2 MerR bends the operator towards itself, whereas Hg-MerR does not significantly bend DNA. *a*, Schematic of the operator constructs phased against a reference bend consisting of six phased A-tracts, with a projected locus between the third and fourth A-tract and an overall bend towards the major groove¹³. The distance between the centre of the operator and the reference bend is 76 bp for the 10-bp linker construct, and consecutive constructs vary by single-base-pair increments. The average helical repeat for this construct is measured to be ~ 10.5 bp per turn¹³. Thus there are ~ 7.5 helical turns between bend centres in a construct with a 13-bp linker. *b*, The mobility of the operator fragments (0.5–3.0 nM) electrophoresed in the presence of MerR (30 nM) with or without 50 μ M Hg(II). The dissociation of the Hg-MerR complexes is consistent with its 3–4-fold lower binding affinity for the operator relative to MerR^{4,8}. Excess Hg(II) under our high thiol conditions does not alter the mobility of the free DNA (not shown). *c*, Plot of the relative electrophoretic migration of phased operator fragments versus the linker length. The absolute migration of the DNA from the well (μ_{abs}) is normalized against the average migration (μ_{av}) of the corresponding set of fragments. Markers represent the normalized values and lines represent the data fitted to a sine function. Filled diamonds represent free DNA; filled circles, MerR; crossed open circles, Hg-MerR; and dashed line, sine fit (using the program IGORpro) of the free DNA. The sinusoidal migration pattern of the free DNA indicates that the left half of these operator constructs is weakly bent. Further normalization against unbound DNA barely alters the phase or the amplitude of the MerR-induced sine wave; however, normalization of low-amplitude sine waves, such as that induced by Hg-MerR, can be misleading in this and in other cases²⁸.

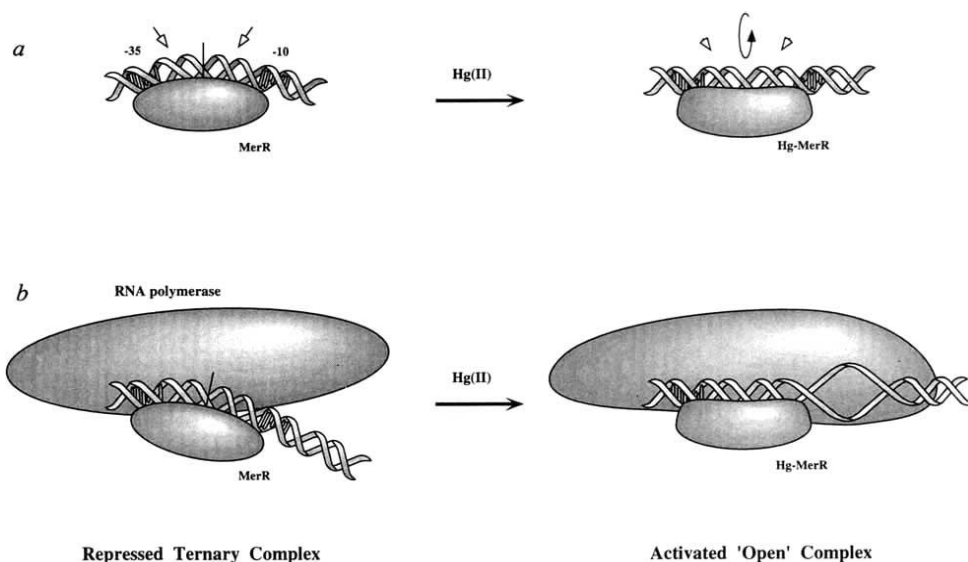


METHODS. A 33-bp *mer* operator was cloned into the *SacI* site of Drak's SB series¹³. The original two A-tract bend of the SB series was removed and single-base linker intermediates were constructed to increase the sensitivity of this system. This new series has a unique *EcoRI* site abutting the operator. *Asel* digestion gives fragments with equidistant arms flanking either the *mer* operator or the reference bend. Electrophoresis of the protein-DNA complexes was carried out as in ref. 5 in 5% 30:1 polyacrylamide on 40-cm gels at 10 $^\circ$ C.

in the CC site show the base step at which CAP induces a kink¹⁹; filled squares in the *mer* operator represent the proposed MerR-induced kink site. The asterisk in the Hg-MerR/DNA complex identifies the site of the Hg(II)-responsive decrease in cleavage on either strand. Published footprinting data on the CAP/DNA site strongly resembles MerR footprints, strengthening the validity of the correction. Furthermore, when a critical Glu residue in the helix-turn-helix motif of each protein (E22 for MerR, E181 for CAP) is mutated to Lys specific DNA binding is abolished²⁹. d, A SETOR display of only the helix-turn-helix motifs of CAP docked in the major groove of the DNA from the coordinates of CAP-DNA crystal structure¹⁹. The red and maroon labelled phosphodiester correspond to the longer and shorter arrows in c, respectively. The CAP-induced ~40° kink site shows a widening of the minor groove to ~10° (ref. 19).

METHODS. A 30–100 nM 3' labelled Styl-Asel fragment of pGMer was incubated with 80 nM MerR ± 3 μM Hg(II) and subjected to DNase I and hydroxyl-radical footprinting^{4,8} and analysed using the program GSXL2-1 (Pharmacia).

FIG. 4 a, Model of MerR-induced kinks. Open arrowheads indicate the location of the kinks, modelled here at CpG steps where the trajectory of the helix axis changes by $\sim 15^\circ$ at each half site. The symmetrical disposition about the centre of the operator (vertical line) is consistent with the projected bend locus (Fig. 1). In the presence of Hg(II), MerR untwists the centre (circular arrow) of the operator^{2,5} and unfixes the operator. Methylation protection and interference data⁴ support the notion that the kinks in each half site operator may arise from docking of the recognition helix of the MerR helix-turn-helix in the major groove at the conserved 5'-GTACG-3' sequence⁹. b, Asymmetric propagation of symmetric DNA distortions. In a ternary complex with a poised RNAP, the polymerase will restrict movement of the -35 half of the operator, the unkinking and untwisting will asymmetrically affect the -10 half of the binding site. The model is supported by three lines of evidence. First, in the absence of MerR, polymerase in closed complexes makes several contacts with the -10 region of the P_T promoter². In the presence of MerR, polymerase does not make these contacts, suggesting that MerR interferes with this interaction (Fig. 3a). Second, polymerase interacts with the minor groove in the -16 region of the promoter³⁰. An MerR-induced kink towards the major groove at -17 directly deflects this region away from the polymerase in the repressor state, whereas Hg(II)-responsive



unkinking brings this region closer to the polymerase in the activator state. Finally, torsional deformation of the spacer is important for open complex formation. MerR bending of the spacer apparently raises the energy required by polymerase to deform the promoter, whereas Hg-MerR untwisting of the spacer^{2,5} lowers the activation barrier, thereby potentiating formation of a transcription bubble and initiation of RNA synthesis.

Separation of the thermodynamic linking number difference into specific alterations in twist and writhe is not feasible in the absence of additional conformational data²³. The bend-phasing data, however, provide useful boundary conditions in the absence of a detailed structure. Given the DNA unbending that accompanies conversion of MerR to Hg-MerR, the ΔLk_3 term need not arise from a simple -33° change in twist as we previously suggested⁵. The activator-induced changes in DNA structure are more easily described for the ΔLk_2 step, where the small mobility differences across the Hg-MerR/DNA and free DNA series suggests minimal bending in the two states. If the ΔLk_2 step involves no change in writhe, Hg-MerR untwists the spacer region of this promoter by $\pm 50^\circ$ relative to free DNA. It is attractive to speculate that upon conversion of the repressor to an activator, a loss of writhe is converted to a decrease in twist, minimizing the thermodynamic cost of distorting the spacer region of P_T .

Although distortions induced by MerR $\pm$ Hg(II) appear symmetrical, MerR allows the RNA polymerase to make an asymmetric ternary complex at P_T both *in vitro* and *in vivo* (Fig. 3a)^{2,3,5}. In this inactive complex, the polymerase makes contacts with only those regions of the promoter that are upstream of -10. Given that the polymerase is ~ 14 times the mass of Hg-MerR, the Hg(II)-responsive unkinking and untwisting of the spacer DNA may displace the unrestrained -10 half of the operator, moving it and the downstream region of the promoter in space. The stereochemistry of the MerR-induced kink towards the major groove at -17 contributes to repression by preventing polymerase from accessing the -10 hexamer, and both bends make it energetically less favourable for RNA polymerase to underwind the spacer DNA. The Hg(II)-responsive unbending, together with the underwinding of the centre of the operator⁵, then spatially repositions the -10 hexamer into the polymerase active site. Underwinding of the spacer may further lower the barrier to formation of the strand-separated open complex. In this regard, MerR functions by modulating promoter structure in a manner that directly assists or inhibits activity of the RNA

polymerase subunit that contacts the -10 and -35 promoter elements, namely σ^{70} . As with many other systems, protein-protein interactions between RNA polymerase subunits and MerR may play a role in regulation²⁴; however, specific roles for modulation of DNA structure in the allosteric mechanism are most apparent.

DNA modulation is also seen in OccR, OxyR and SoxR (refs 25–27, respectively) and other members of the LysR and MerR families of ligand-responsive transcriptional regulators. As studies of eukaryotic transcriptional mechanisms reveal poised multifactor arrays that can activate only in a specific spatial context¹, and as additional regulatory factors lacking activation domains are found, it is likely that the significance of protein-induced modulation of DNA structure in transcriptional regulation will become apparent⁹.

Note added in proof: It has recently been shown³¹ that the widened minor grooves characteristic of the CAP-induced bends are preferred sites for both DNase I and IN integrase activity. $\square$

Received 25 October 1994; accepted 10 February 1995.

1. Tjian, R. & Maniatis, T. *Cell* **77**, 5–8 (1994).
2. Frantz, B. & O'Halloran, T. V. *Biochemistry* **29**, 4747–4751 (1990).
3. Heltzel, A., Lee, I. W., Totis, P. A. & Summers, A. O. *Biochemistry* **29**, 9572–9584 (1990).
4. O'Halloran, T. V., Frantz, B., Shin, M. K., Ralston, D. M. & Wright, J. G. *Cell* **56**, 119–129 (1989).
5. Ansari, A. Z., Chael, M. L. & O'Halloran, T. V. *Nature* **355**, 87–89 (1992).
6. Comess, K. M., Shewchuk, L. M., Ivanetich, K. & Walsh, C. T. *Biochemistry* **33**, 4175–4186 (1994).
7. Parkhill, J. & Brown, N. L. *Nucleic Acids Res.* **18**, 5157–5162 (1990).
8. Parkhill, J., Ansari, A. Z., Wright, J., Brown, N. L. & O'Halloran, T. V. *EMBO J.* **12**, 413–421 (1993).
9. Ansari, A. Z. & O'Halloran, T. V. in *Transcription: Mechanisms and Regulation* (eds Conaway, R. C. & Conaway, J. W.) 369–386 (Raven, New York, 1994).
10. Zinkel, S. S. & Crothers, D. M. *Nature* **328**, 178–181 (1987).
11. Salvo, J. J. & Grindley, N. D. F. *Nucleic Acids Res.* **15**, 9771–9779 (1987).
12. Crothers, D. M., Gartenberg, M. R. & Shrader, T. E. *Meth. Enzym.* **208**, 118–146 (1991).
13. Drak, J. & Crothers, D. M. *Proc. natn. Acad. Sci. U.S.A.* **88**, 3074–3078 (1991).
14. Kahn, J. D., Yun, E. & Crothers, D. M. *Nature* **368**, 163–166 (1994).
15. Lahm, A., Weston, S. A. & Suck, D. *Nucleic Acids molec. Biol.* **5**, 171–186 (1991).
16. Hogan, M. E., Roberson, M. W. & Austin, R. H. *Proc. natn. Acad. Sci. U.S.A.* **86**, 9273–9277 (1989).
17. Travers, A. A. & Klug, A. in *DNA Topology an its Biological Effects* (eds Cozzarelli, N. R. & Wang, J. C.) 57–106 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1990).

18. Gaston, K., Bell, A., Kolb, A., Buc, H. & Busby, S. *Cell* **62**, 733–743 (1990).
19. Schultz, S. C., Shields, G. C. & Steitz, T. A. *Science* **253**, 1001–1007 (1991).
20. Raumann, B. E., Rould, M. A., Pabo, C. O. & Sauer, R. T. *Nature* **367**, 754–757 (1994).
21. Burkhoff, A. M. & Tullius, T. D. *Nature* **331**, 455–457 (1989).
22. Hunter, C. A. *J. molec. Biol.* **256**, 1025–1054 (1993).
23. Cozzarelli, N. R., Boles, T. C. & White, J. H. in *DNA Topology and its Biological Effects* (eds Cozzarelli, N. R. & Wang, J. C.) 139–184 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1990).
24. Ptashne, M. in *A Genetic Switch* 13–123 (Cell press & Blackwell Scientific, Cambridge, MA, 1986).
25. Wang, L., Helmann, J. D. & Winans, S. C. *Cell* **69**, 659–667 (1992).
26. Storz, G., Tartaglia, L. A. & Ames, B. N. *Science* **248**, 189–194 (1990).
27. Hidalgo, E. & Demple, B. *EMBO J.* **13**, 138–146 (1993).
28. Snyder, U. K., Thompson, J. F. & Landy, A. *Nature* **341**, 255–257 (1989).
29. Ross, W., Park, S.-J. & Summers, A. O. *J. Bact.* **171**, 4009–4018 (1989).
30. Borowicz, J. A. & Gralla, J. D. *Biochemistry* **25**, 5051–5057 (1986).
31. Müller, H.-P. & Varmus, H. E. *EMBO J.* **13**, 4704–4714 (1994).

ACKNOWLEDGEMENTS. We thank J. Kahn and D. M. Crothers for phasing plasmids; J. G. Wright for MerR protein; J. Jarzembowski, S. Fox, C. Lima and J. Roca for assistance; J. Kahn, M. Chael, M. Grieserstein, C. Cera, M. Shida and members of the O'Halloran group for discussion; and the NIH and A. P. Sloan Foundation for funding.

Structure of a multisubunit complex that promotes DNA branch migration

Carol A. Parsons, Andrzej Stasiak*,
Richard J. Bennett & Stephen C. West†

Genetic Recombination Laboratory, Imperial Cancer Research Fund, Clare Hall Laboratories, South Mimms, Hertfordshire EN6 3LD, UK
* Laboratory of Ultrastructural Analysis, University of Lausanne, 1015 Lausanne, Switzerland

THE RuvA and RuvB proteins of *Escherichia coli*, which are induced in response to DNA damage, are important in the formation of heteroduplex DNA during genetic recombination and related recombinational repair processes¹. *In vitro* studies show that RuvA binds Holliday junctions^{2–5} and acts as a specificity factor that targets the RuvB ATPase, a hexameric ring protein^{6,7}, to the junction. Together, RuvA and RuvB promote branch migration, an ATP-dependent reaction that increases the length of the heteroduplex DNA^{3,8–10}. Electron microscopic visualization of RuvAB now provides a new insight into the mechanism of this process. We observe the formation of a tripartite protein complex in which RuvA binds the crossover and is sandwiched between two hexameric rings of RuvB. The Holliday junction within this complex adopts a square-planar structure. We propose a molecular model for branch migration, a unique feature of which is the role played by the two oppositely oriented RuvB ring motors.

To visualize the interaction of RuvA with Holliday junctions, we prepared χ -structures in which a Holliday junction was flanked by four duplex arms with lengths of about 660, 850, 1,500 and 1,940 base pairs (bp). Nitrocellulose filter binding assays demonstrated the specific formation of RuvA- χ -structure complexes which formed in the absence (Fig. 1a) or presence (Fig. 1b) of divalent metal ions (Mg^{2+}). Although more protein was required in the presence of Mg^{2+} , in both cases the specificity of RuvA for χ -DNA was higher than for linear duplex DNA.

To visualize the binding of RuvA to the χ -structure, conditions were used to favour the most stable association of RuvA with DNA. Binding complexes, prepared in the absence of metal ions, were visualized by electron microscopy after negative staining with uranyl acetate. The specificity of RuvA for the crossover was apparent and binding to the flanking duplex DNA was not observed (Fig. 1d, e). The RuvA protein (which is tetrameric in solution)^{6,11} appeared as homogeneous 8.5–10-nm particles characteristic of a RuvA multimer. Similar complexes were observed in the presence of Mg^{2+} (data not shown).

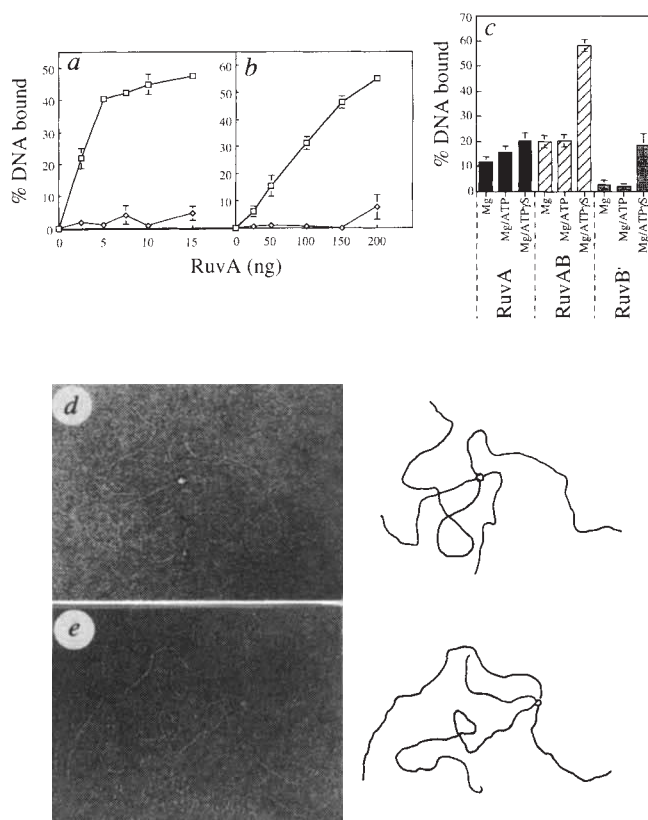


FIG. 1 Specific binding of χ -structures by RuvAB proteins. a, b, Nitrocellulose filter assay showing the binding of χ -DNA ($\square$) or linear duplex DNA ($\diamond$) by RuvA in the absence (a) or presence (b) of divalent ions. c, Filter assay indicating the binding of RuvA, RuvAB or RuvB to χ -DNA under a variety of reaction conditions. d, e, Visualization of RuvA-Holliday junction complexes by electron microscopy.

METHODS. RuvA and RuvB were prepared as described¹¹. Plasmid pSD115 was grown in *E. coli* strain RM40 as described²². RM40 contains the *xerC* gene under control of the *lac* promoter. pSD115 contains a duplication of the *cer* recombination site which recombines in response to expression of XerC (induction with IPTG leads to the formation of figure-eight DNA molecules). Plasmid DNA was prepared from non-induced or induced cultures to generate form I or figure-eight DNAs, respectively. Linear duplex DNA was made by restriction cleavage of form I pSD115 with *StyI*. χ -DNA was made by treating figure-eight DNA with *Scal* and *StyI* followed by sucrose gradient centrifugation. In a, binding reactions (50 μ l) contained 3'-³²P end-labelled χ -DNA or linear duplex DNA (50 ng) in TED buffer (20 mM triethanolamine-HCl, pH 7.5, 5 mM EDTA and 1 mM dithiothreitol) with RuvA protein as indicated. For the reactions in b, EDTA was replaced with 10 mM $MgCl_2$ (TMD buffer). For c, reactions were done in TMD plus ATP (0.5 mM) or ATP- γ S (0.25 mM). RuvA (50 ng) and/or RuvB (100 ng) were present as indicated. Incubation was for 15 min at 37 °C followed by fixation by incubation with 0.2% glutaraldehyde at 37 °C for 15 min. For the filter assay, reactions were supplemented with 450 μ l of the appropriate iced binding buffer, applied to nitrocellulose filters (Schleicher & Schuell, NC45) and washed with 2 ml of iced buffer. Filters were dried and the per cent ³²P label retained was determined using a phosphorimager. Background binding in the absence of enzyme has been subtracted. Each data point represents the average of several experiments. Standard error bars are shown. For electron microscopy (d and e), binding reactions (20 μ l) contained χ -DNA (50 ng) and RuvA (10 ng) in TED buffer. Incubation was for 15 min at 37 °C and complexes were fixed with glutaraldehyde. For spreading and adsorption, DNA-protein complexes were diluted to a concentration of 1 μ g ml⁻¹ in 5 mM magnesium acetate pH 7.0. RuvA was visualized after staining with 2% uranyl acetate. Images were recorded under minimal dose conditions at $\times 35,000$ on a Philips CM12 electron microscope. Our interpretation of each image is indicated on the right. Scale bar, 100 nm.

† To whom correspondence should be addressed.

Nitrocellulose filter binding assays were used to determine the conditions required for the formation of RuvAB–Holliday junction complexes (Fig. 1c). The presence of Mg^{2+} and ATP- γ S (a non-hydrolysable analogue of ATP), favoured the formation of stable complexes without branch migration³. Binding reactions containing χ -DNA, RuvA and RuvB were spread and visualized by electron microscopy following negative staining. Well defined protein complexes were observed at the site of the crossover, and a region of a specimen containing three such complexes is shown in Fig. 2. Each complex has a tripartite structure in which the central particle resembles the RuvA oligomer, and the two flanking particles exhibit the shape and dimensions characteristic of a hexameric RuvB ring⁷. Preliminary image analysis of the tripartite complexes indicates that the two RuvB rings have a mirror symmetry, with their more convex surfaces facing outwards (E. Egelman, personal communication). Double hexameric rings of RuvB assembled on duplex DNA show a similar structural symmetry⁷. In the absence of RuvA, tripartite complexes were not formed. When similar reactions were done using ATP rather than ATP- γ S, and specimens were fixed with glutaraldehyde, tripartite structures were again observed by microscopy.

Using 50 ng RuvA and 100 ng RuvB, we observed that most of the RuvAB complexes were tripartite (69%), but also observed more complex structures in which RuvA was flanked by three (27%) or four (4%) rings of RuvB (data not shown). Because DNase I protection studies of complexes formed between RuvAB and small synthetic Holliday junctions demonstrate that the binding of only two arms of DNA by RuvB is sufficient for branch migration⁵, the tripartite complexes are likely to represent the active species of the RuvAB/branch migration complex.

Electron microscopic observations of the tripartite RuvAB structure showed that the rings of RuvB were diametrically opposed, indicating that the DNA had adopted a form characteristic of an unfolded square-planar structure. To determine whether the structure of the Holliday junction was directed by its interaction with RuvA, the structure of the RuvA–Holliday junction complex was determined by gel electrophoresis. For this assay, a small synthetic Holliday junction was prepared by annealing four oligonucleotides. The sequences of the DNA strands were arranged so that each arm contained a unique

*Bam*HI (B), *Hind*III (H), *Eco*RI (R) or *Xba*I (X) restriction site (Fig. 3a). Comparison of the electrophoretic mobility of the six species resulting from cleavage with two restriction enzymes (to produce molecules with two long arms and two short arms) allows a determination of junction structure^{12,13}.

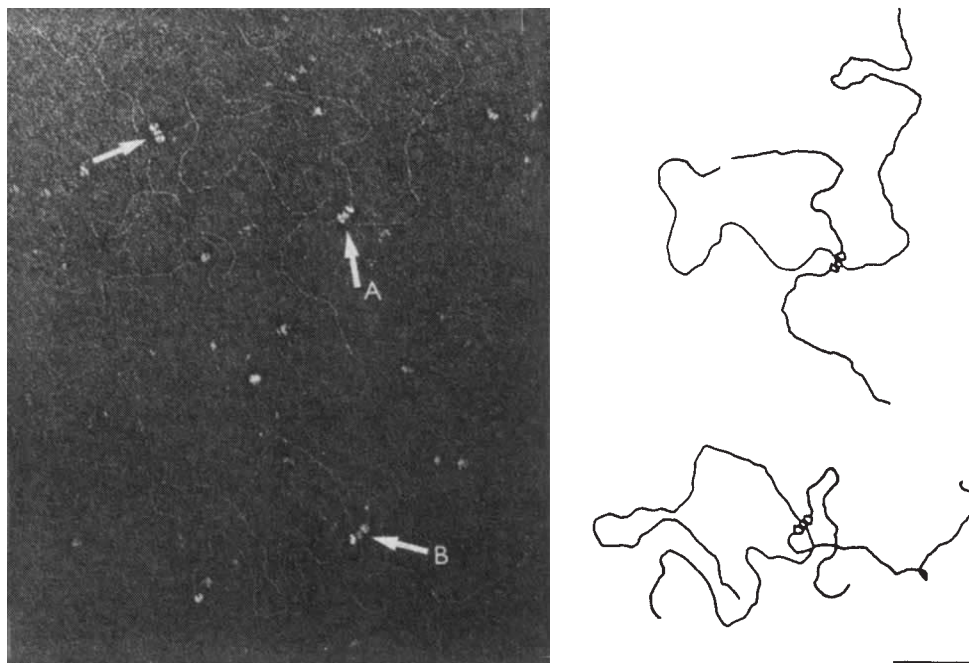
In control experiments, the structure of the naked DNA junction was determined. In the absence of Mg^{2+} , a gel pattern corresponding to four slow and two fast species (4:2 pattern) was observed, characteristic of an extended structure in which the four arms of DNA extend towards the corners of a square (Fig. 3b). This unfolded structure is defined as the square-planar form of the junction¹⁴. In contrast, in the presence of Mg^{2+} ions (Fig. 3c), the junction had an electrophoretic mobility characteristic of a folded structure (stacked X-structure) with three pairs of comigrating species (2:2:2 pattern). These results agree with previous observations^{12,14}.

When RuvA–Holliday junction complexes were formed in the presence of Mg^{2+} and analysed using the same electrophoretic mobility technique, the resulting band-shift complexes showed a 4:2 pattern characteristic of the square-planar junction configuration (Fig. 3d). These results show that the binding of RuvA either results in an unfolding of the Holliday junction or a stabilization of the unfolded state of the junction.

The low-resolution structure of the RuvAB–Holliday junction complex determined by electron microscopy leads us to propose a new, yet simple mechanism for the formation of heteroduplex DNA (Fig. 4a). There are two key features to this model. First, the Holliday junction is unfolded and lies in a square-planar configuration. Second, branch migration is catalysed by a tripartite protein complex in which two hexameric RuvB rings sandwich RuvA. The two ring motors, which are diametrically opposed across the Holliday junction, each encompass a DNA duplex and drive branch migration by catalysing the directional unwinding and helical rotation of DNA^{15,16}.

As this study indicates, the Holliday junction is unfolded by RuvA. Previous studies have shown that the rate of spontaneous branch migration of an unfolded (protein-free) Holliday junction is about 1,000 times faster than that observed with a folded junction^{17,18}. Because the isomeric form of a folded junction (the choice of which arms stack as partners) is determined by the local DNA sequence at the junction point¹², branch migration is energetically unfavourable because the junction may have to

FIG. 2 Visualization of tripartite RuvAB–Holliday junction complexes by electron microscopy. Three specific complexes can be observed in this section of the grid, and our interpretations of two molecules (A and B) are shown to the right. Binding reactions (20 μ l) contained χ -DNA (50 ng), RuvA (50 ng) and RuvB (100 ng) in TMD buffer + 0.25 mM ATP- γ S. Incubation was for 15 min at 37 °C. EM was as in Fig. 1 legend.



switch isomeric form at each step. This would require unstacking followed by restacking, requiring large relative movements of the helical arms at the junction point¹⁴. In contrast, branch migration of an unfolded junction would be more efficient because it requires only the axial rotation of DNA, without need for any isomeric transitions.

Although branch migration between two homologous DNA molecules is isoenergetic, nucleotide binding and hydrolysis by RuvAB are required for branch migration. We propose that the reaction involves ATP-dependent unwinding and helical rotation of DNA (Fig. 4a), which causes the DNA to move with a defined polarity through each RuvB ring. Because the two RuvB ring motors have opposite orientations, the DNA may pass through the branch migration machine as shown in Fig. 4b. The consequence of such a mechanism is the conversion of homoduplex DNA, which enters the RuvAB complex, into heteroduplex DNA which exits after strand exchange. Because of the symmetry of a Holliday junction, RuvAB-mediated branch migration can occur in either direction^{2,8,9}, depending upon which pair of arms are bound by RuvB⁵. Consistent with recent DNase I footprinting studies⁵, directionality has been incorporated into

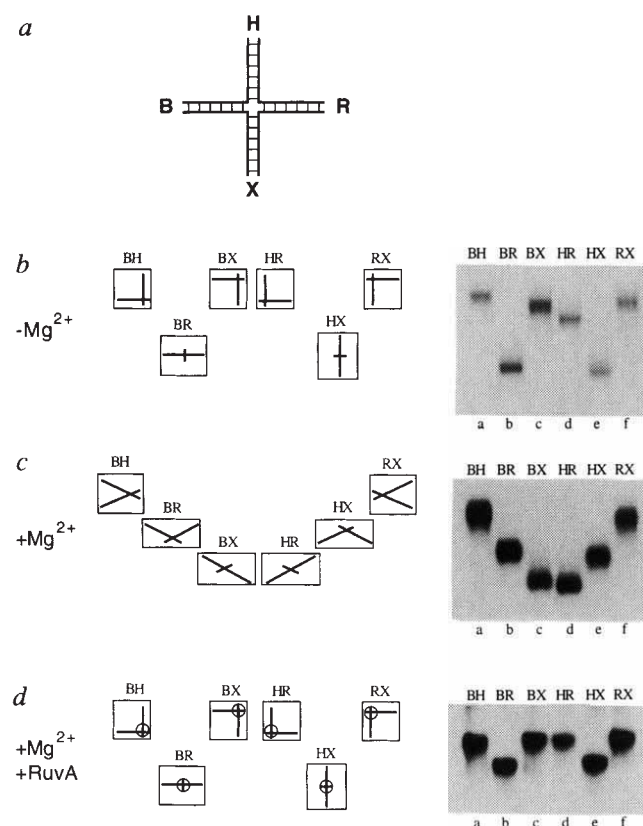


FIG. 3 Unfolding of a Holliday junction by RuvA protein. **a**, Diagram of the synthetic Holliday junction containing four restriction sites: *Bam*HI, *B*, *Hind*III, *H*, *Eco*RI, *R* and *Xba*I, *X*. The construction, 5' ³²P-end labelling, and preparation of the Holliday junction (junction 3) from four 80-mer oligonucleotides was essentially as described¹². It was digested with pairs of restriction enzymes to generate six junctions with two long and two short arms. Labels above each lane indicate which arms of the junction are left intact after restriction digestion. **b–d**, ³²P-labelled junctions (1 ng) were incubated alone or with RuvA (100 ng) in 50 mM Tris-HCl pH 8.0, 1 mM dithiothreitol and 100 µg ml⁻¹ BSA. EDTA (5 mM) or MgCl₂ (0.5 mM) were present as indicated. Incubation was for 15 min at 20 °C. Samples were loaded onto 8% polyacrylamide gels and electrophoresed at 150 V for 15 h (no protein) or at 200 V for 24 h (in presence of RuvA). The gel and the buffer system contained 89 mM Tris-borate and either (b) 2 mM EDTA (–Mg²⁺) or (c and d) 0.5 mM MgCl₂ (+Mg²⁺). Gels were dried and the DNA visualized by autoradiography on Kodak XAR film.

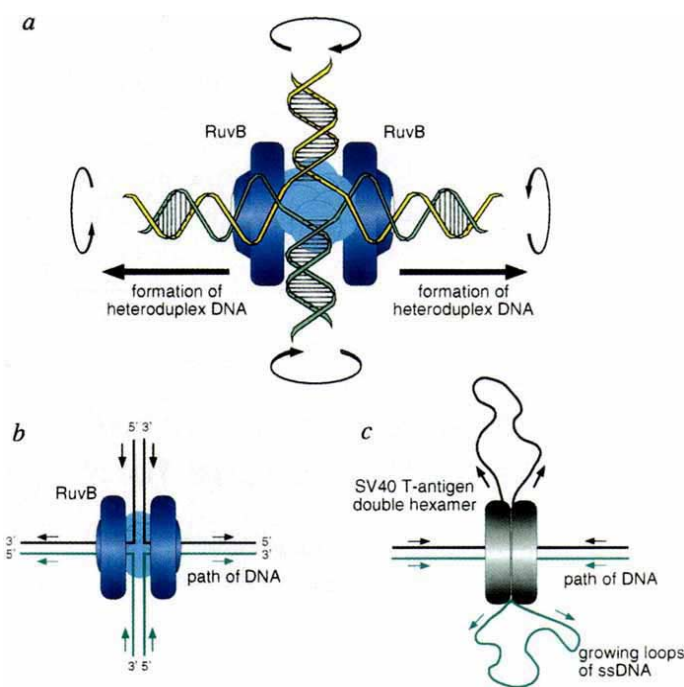


FIG. 4 Model for RuvAB-mediated branch migration. **a**, Diagram indicating the structure of the RuvAB-Holliday junction complex in which RuvA binds the crossover and is sandwiched between two hexameric rings of RuvB. Each ring encompasses a DNA duplex (in this cut-away view only half of each ring is shown). The DNA lies in a square planar configuration. Branch migration occurs as the RuvB ring motors, which lie in opposite orientations, promote the passage of DNA. Further details of the mechanism of branch migration, which may involve strand separation, are described in the text. **b** and **c**, Diagrams indicating the similarity between branch migration and unwinding of a replication origin catalysed by the hexameric rings of RuvB and SV40 T-antigen, respectively. The directionality of strand passage differs in each case.

the model shown in Fig. 4b. With fully homologous DNAs, passage through the RuvAB complex will lead to branch migration where every base pair opened in the substrate will be compensated by a newly formed base pair in the products. However, when the DNA substrates contain heterologies, the heteroduplex products would contain mismatches or loops that would remain unpaired. The ATP-dependent DNA helicase activity of RuvAB may aid branch migration through heterologies^{2,19}.

Characterization of the *E. coli* RuvAB branch migration complex provides direct insight into the mechanism of a molecular motor which functions in the formation of heteroduplex DNA during genetic recombination and the recombinational repair of DNA damage. Remarkably, this mechanism of branch migration shares a number of similarities with the way in which simian virus 40 (SV40) large T-antigen initiates DNA replication because electron microscopic observations indicate that T-antigen forms a double hexameric ring structure at the SV40 origin²⁰. In this case, the two T-antigen rings catalyse the passage of duplex DNA into the complex, where it is unwound to form single-stranded DNA loops²¹ (Fig. 4c). The passage of DNA through a pair of ring proteins that are targeted to a specific DNA structure or sequence, suggests a new and efficient mechanism for DNA movement that may be used in other aspects of DNA metabolism. □

Received 23 December 1994; accepted 8 February 1995.

- West, S. C. *Cell* **76**, 9–15 (1994).
- Parsons, C. A., Tsaneva, I., Lloyd, R. G. & West, S. C. *Proc. natn. Acad. Sci. U.S.A.* **89**, 5452–5456 (1992).
- Parsons, C. A. & West, S. C. *J. molec. Biol.* **232**, 397–405 (1993).
- Iwasaki, H., Takahagi, M., Nakata, A. & Shinagawa, H. *Genes Dev.* **6**, 2214–2220 (1992).

5. Hiom, K. & West, S. C. *Cell* **80**, 787–793 (1995).
6. Mitchell, A. H. & West, S. C. *J. molec. Biol.* **243**, 208–215 (1994).
7. Stasiak, A. et al. *Proc. natn. Acad. Sci. U.S.A.* **91**, 7618–7622 (1994).
8. Tsaneva, I. R., Müller, B. & West, S. C. *Cell* **69**, 1171–1180 (1992).
9. Müller, B., Tsaneva, I. R. & West, S. C. *J. biol. Chem.* **268**, 17179–17184 (1993).
10. Shiba, T., Iwasaki, H., Nakata, A. & Shinagawa, H. *Proc. natn. Acad. Sci. U.S.A.* **88**, 8445–8449 (1991).
11. Tsaneva, I. R., Illing, G. T., Lloyd, R. G. & West, S. C. *Molec. gen. Genet.* **235**, 1–10 (1992).
12. Duckett, D. R. et al. *Cell* **55**, 79–89 (1988).
13. Duckett, D. R., Murchie, A. I. H. & Lilley, D. M. J. *EMBO J.* **9**, 583–590 (1990).
14. Lilley, D. M. J. & Clegg, R. M. A. *Rev. Biophys. biomolec. Struct.* **22**, 299–328 (1993).
15. Tsaneva, I. R., Müller, B. & West, S. C. *Proc. natn. Acad. Sci. U.S.A.* **90**, 1315–1319 (1993).
16. Adams, D. E. & West, S. C. *J. molec. Biol.* **247**, 404–417 (1995).
17. Müller, B., Burdett, I. & West, S. C. *EMBO J.* **11**, 2685–2693 (1992).
18. Panyutin, I. G. & Hsieh, P. *Proc. natn. Acad. Sci. U.S.A.* **91**, 2021–2025 (1994).
19. Iype, L. E., Wood, E. A., Inman, R. B. & Cox, M. M. *J. biol. Chem.* **269**, 24967–24978 (1994).
20. Dean, F. B., Borowiec, J. A., Eki, T. & Hurwitz, J. *J. biol. Chem.* **267**, 14129–14137 (1992).
21. Wessel, R., Schwizer, J. & Stahl, H. *J. Virol.* **66**, 804–815 (1992).
22. McCulloch, R., Coggins, L. W., Colloms, S. D. & Sherratt, D. J. *EMBO J.* **13**, 1844–1855 (1994).

ACKNOWLEDGEMENTS. We thank D. Adams, E. Egelman, N. Hajibagheri, K. Hiom, A. Mitchell and I. Tsaneva for discussions and communication of unpublished data. A. Mitchell and I. Tsaneva supplied RuvA and RuvB and D. Sherratt provided pSD115. This work was supported by the ICRF, the Swiss National Foundation and the British-Swiss Joint Research Program.

Crystal structure of the nuclear Ras-related protein Ran in its GDP-bound form

Klaus Scheffzek*, Christian Klebe*, Karin Fritz-Wolf†, Wolfgang Kabsch† & Alfred Wittinghofer*‡

* Max-Planck-Institut für molekulare Physiologie, Abteilung Strukturelle Biologie, Rheinlanddamm 201, 44139 Dortmund, Germany

† Max-Planck-Institut für medizinische Forschung, Abteilung Biophysik, Jahnstrasse 29, 69120 Heidelberg, Germany

‡ To whom correspondence should be addressed

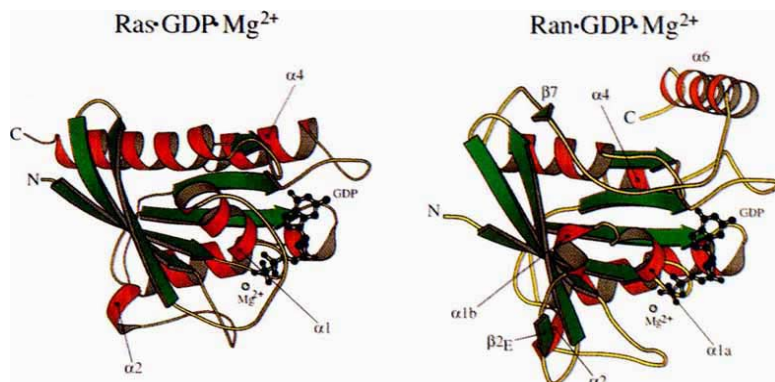
THE Ran proteins constitute a distinct branch of the superfamily of Ras-related GTP-binding proteins¹ which function as molecular switches cycling between GTP-bound 'on' and GDP-bound 'off' states². Ran is located predominantly in the nucleus of eukaryotic cells³ and is involved in the nuclear import of proteins^{4,5} as well as in control of DNA synthesis and of cell-cycle progression^{6–8}. We report here the crystal structure at 2.3 Å resolution of human Ran (*M_r* 24K) complexed with GDP and Mg²⁺. This structure reveals a similarity with the Ras core (G-domain) but with significant variations in regions involved in GDP and Mg²⁺ coordination (switch I and switch II regions in Ras)^{9,10}, suggesting that there could be major conformational changes upon GTP binding. In addition to the G-domain, an extended chain and an α -helix were identified at the carboxy terminus. The amino-terminal (amino-acid residues ¹MAAQGE⁷) stretch and the acidic tail (²¹¹DEDDDL²¹⁶) appear to be flexible in the crystal structure.

The structure was solved with the multiple isomorphous replacement method (MIR; Table 1). The present model comprises residues 8 to 210 (out of 216 in full-length Ran), GDP, Mg²⁺ and 25 water molecules. A ribbon plot with the secondary structure elements of the Ran model is shown in Fig. 1 and compared with the Ras model. The main features of the G-domain, consisting of a six-stranded β -sheet surrounded by 5 α -helices connected with loops, as found in Ras^{9–11}, EF-Tu^{12–14}, transducin- α ¹⁵, G α_i ¹⁶ and EF-G^{17,18}, are basically conserved; however, Ran shows significant deviations within the G-domain.

Helix α_1 has a kink after the second turn which results in a completely different orientation of what is called the effector region in Ras. In Ran, residues 38–40 of this region form a short β -strand (β_{2E} , or 'effector' β -strand) antiparallel to β_2 . Unlike the situation with Ras, helix α_2 and loop L4 are well defined in the Ran-GDP structure^{9,11}. Helix α_4 , which contains five residues, is considerably shorter and, together with five preceding residues, looks like a distorted version of its Ras equivalent. This possibly arises from the lack of one residue in the Ran sequence, as can be seen from the sequence and secondary-structure alignment (Fig. 2).

In addition to the G-domain, we found a short β -strand, β_7 (antiparallel to β_2), a long extended chain containing a number of prolines and a five-turn α -helical extension, helix α_6 , at the C terminus, which appears to be detached from the rest of the protein. It has comparable orientation and appears to be stabilized by similar intramolecular interactions in monoclinic and tetragonal Ran crystals. In addition, it is involved in crystal contacts which are different for the two crystal forms (details will be presented elsewhere). The acidic tail (²¹¹DEDDDL²¹⁶) seems to be flexible, as judged from poor electron density, high B-factors, and from the observation that the preceding residues (208–210) assume different conformations in two crystal forms.

FIG. 1. Ribbon representation of Ran-GDP compared with Ras-GDP²⁶. The current model and the model of Ras-GDP (from the Brookhaven Data Bank; 1q21) were analysed by the program DSSP²⁹ and drawn using the program Molscript³⁰. Structural comparison of Ras and Ran was carried out with the Isq-explicit option of program O (ref. 27). Secondary-structure elements that differ between Ras and Ran are indicated (see text; also Fig. 2).



The binding mode of nucleotide and Mg^{2+} shows many features characteristic of GTP-binding proteins in general and some specific for Ras-related proteins (Fig. 3). The guanine base interacts in the usual manner with lysine (K123) and aspartic acid (D125) from the highly conserved NKXD motif. Unlike Ras, in which F28 (stabilized by K147) interacts hydrophobically

with the base, the corresponding residue in Ran (F35) is found before $\beta 2_E$, 25 Å away from the C5 atom of guanine. It appears that the aromatic interaction stabilizing guanine binding in Ras is substituted by the side chain of K152 (K147 in Ras) which is oriented parallel to K123 on the opposite side of the purine ring. The 2' and 3' hydroxyl groups of the ribose are exposed to the

TABLE 1 Data collection and refinement statistics

Crystal form	Native	Tetragonal ($P4_12_12$) 2 mM $K_2Hg(SCN)_4$	5 mM K_2PtBr_4	Monoclinic ($P2_1$) Native
Soaking time (h)		16	24	
Data collection				
Cell dimensions (Å)	$a=b=62.6$ $c=122.3$	$a=b=62.7$ $c=121.5$	$a=b=62.9$ $c=122.1$	$a=59.3$, $b=59.9$ $c=65.6$, $\beta=95.9^\circ$
Molecules/asym.unit (no.)	1			2
Resolution (Å)	2.3	2.9	2.7	2.6
Completeness (%)	96	93	98	92
Obs. reflections (no.)	105,733	28,442	34,390	23,612
Unique reflections (no.)	10,982	5,804	7,116	12,918
R_{sym}^* (%)	6.8	8.2	6.2	6.5
Structure determination		MIR		Molecular replacement
Resolution (Å)	15–3	15–4	15–3	15–4
$R_F^{\dagger}$ (%)		31.7	27.8	
No. of sites		4	3	
$\langle m \rangle^{\ddagger}$	0.44			
$R_C^{\S}$		52.2	54.8	
$F_H/E^{\parallel}$		1.7	1.5	
Rotation search ¶ Θ_1 , Θ_2 , Θ_3				
Molecule A				7.3, 81.6, 4.6
Molecule B				87.1, 80.1, 177.9
Translation search $^{\#}$ x , y , z				
Molecule A				0.29, 0.0, 0.0
Molecule B				0.59, 0.22, 0.92
Final refinement				
Resolution (Å)	10–2.3			10–2.6
R_{cryst} (%) *	22.8			22.7
R_{free} (%) *	29.9			32.1
Waters (no.)	25			16 + 16
R.m.s. bond length deviations (Å)	0.009			0.009
R.m.s. bond angle deviations ($^\circ$)	1.175			1.454

Recombinant full-length human Ran protein was expressed from *E. coli* using the TC4 cDNA described and supplied by P. D'Eustachio and purified as a 1:1 complex with GDP¹⁹. Crystals were grown at room temperature using the hanging-drop method with 500 μ l reservoir and drops containing equal volumes (usually 3 + 3 μ l) of protein (15 mg ml⁻¹, dialysed against water with 12 mM dithiothreitol and reservoir solution (100 mM Tris-HCl, pH 7.5, 20 mM $MgCl_2$, 17–20% PEG3350; Sigma). Under identical conditions, tetragonal and monoclinic crystals were obtained with one and two molecules of Ran in the asymmetric unit. Forty heavy-atom derivatives were prepared by soaking tetragonal crystals for 5–40 h in 300 μ l of reservoir solution containing 0.2–5 mM of the respective heavy-atom compound. X-ray diffraction data were collected with a Siemens/Nicolet area detector using focused X-rays from a GX-18 rotating-Cu-anode generator and were reduced with XDS²⁵. Derivative data were analysed with a program package written by W.K. (unpublished). A 5-Å MIR map became interpretable after solvent flattening and was used for identifying the expected arrangement of α -helices of the G-domain. p21^{ras} α coordinates (residues 1–171 of human H-Ras)²⁶ were positioned in this 5-Å MIR map and served as a frame of reference for model building based on a solvent-flattened 3-Å MIR map. After 20 rounds of successive model building using program O (ref. 27) and simulated annealing refinement with X-PLOR 3.1 (ref. 28), during which the resolution was gradually increased up to 2.3 Å, the R-factor was 26%, including residues 8–210 and GDP. A spherical electron density peak close to the β -phosphate was assigned as Mg^{2+} , which appears to be justified in view of the resulting coordination pattern of the Mg^{2+} ion and its conservation in the high-resolution structures of p21^{ras} (refs 9, 11, 26), EF-Tu (refs 12–14), G_{ai} (ref. 16) and G_{at} (ref. 15). Finally water molecules were added. At present we do not see well defined density for seven N- and six C-terminal residues. The position of the Ran molecule in the tetragonal cell is consistent with a previous molecular replacement solution obtained with X-PLOR 3.1 using the coordinates of p21^{ras} in the GDP-bound form²⁶ as a search model. However, we were not able to build the Ran model from this solution. The model we obtained from solving the tetragonal crystal structure of Ran was used for a molecular replacement search with X-PLOR 3.1 in the monoclinic crystal form, which resulted in two clear solutions for the rotation and translation functions. The structure of the monoclinic Ran was refined with X-PLOR, imposing restrained crystallographic symmetry on residues not involved in crystal contacts. The two independent molecules are related by an approximate twofold rotation. Most of the crystal contacts are different in the two crystal forms. Coordinates of the 2.3 Å Ran model are being submitted to the Brookhaven Data Bank.

* $R_{sym} = \sum |I_h - \langle I \rangle_h| / \sum \langle I \rangle_h$; I , where h is the scaled intensity of the i th symmetry-related observation of reflection h and $\langle I \rangle_h$ is the mean value.

$\dagger R_F = 2 \sum_h |F_{PH} - F_P| / \sum_h |F_P + F_{PH}|$, where F_{PH} and F_P are the derivative and native structure amplitudes, respectively.

$\ddagger \langle m \rangle$, Mean figure of merit.

$\S R_C = \sum |F_{PH,obs} - F_{PH,calc}| / \sum |F_{PH,obs} - F_P|$

$\parallel F_H/E = \sqrt{\sum f_H^2} / \sqrt{\sum (F_{PH,obs} - F_{PH,calc})^2}$

$\P$ Eulerian angles Θ_1 , Θ_2 , Θ_3 as defined in X-PLOR.

$\#$ Fractional coordinates x , y , z as obtained by X-PLOR.

$\star R_{cryst} = \sum |F_{oh} - F_{ch}| / \sum F_{oh}$, where F_{oh} and F_{ch} are the observed and calculated structure factor amplitudes for reflection h . 10% of the reflections were put aside for the calculation of the free R-factor and were not included in the refinement.

solvent, which explains why binding of GDP analogues containing a fluorescent group is not impaired¹⁹.

Details of the binding of the Mg^{2+} cofactor are shown in Fig. 3. A new and unexpected finding is the water-mediated coordination of E70 (E62 in Ras) and T66 (T58 in Ras) to the metal ion. Particularly in the case of E70, the coordination is

unlike the situation in the Ras structure, where the corresponding residue (E62) (together with seven neighbouring residues in the sequence) has been found to be flexible in the crystal structure as well as in the structure determined by NMR. T66 and E70 belong to the totally conserved DTAGQE motif in Ras-related proteins¹. As this stretch is part of the switch II region

FIG. 2 Alignment and secondary-structure assignment of Ras and Ran. The alignment is based on the structure of truncated human H-Ras (residues 1–171) in the GDP-bound form²⁶ as analysed by DSSP²⁹, and that of full-length human Ran presented here. To preserve the secondary-structure nomenclature of the G-domain, the additional β -strand before $\beta 2$ was labelled the effector β -strand $\beta 2_E$, and helix $\alpha 1$ was considered to consist of two parts, $\alpha 1a$ and $\alpha 1b$. The N- and C-terminal amino acids not included in this model are indicated by a wavy line. Residues conserved in all Ras-related proteins¹ are shown in bold.

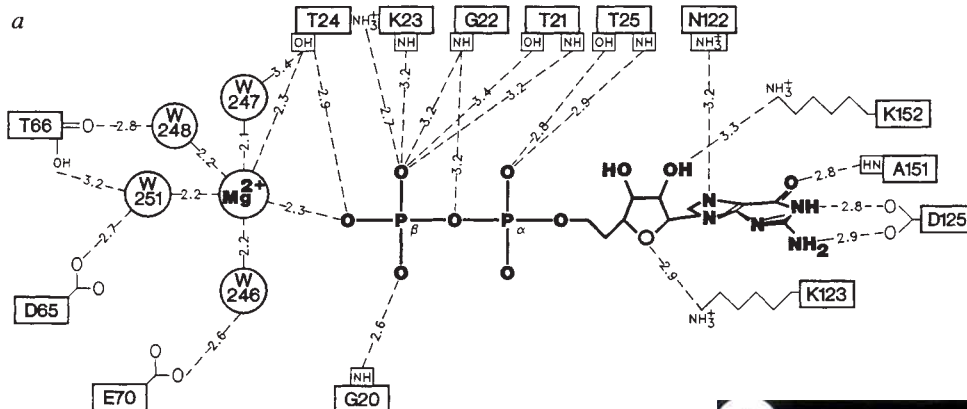
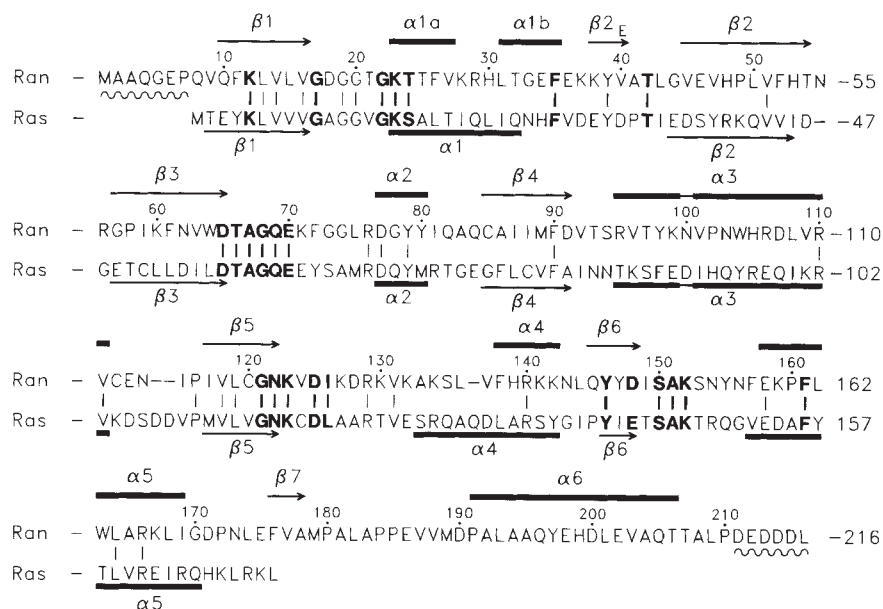
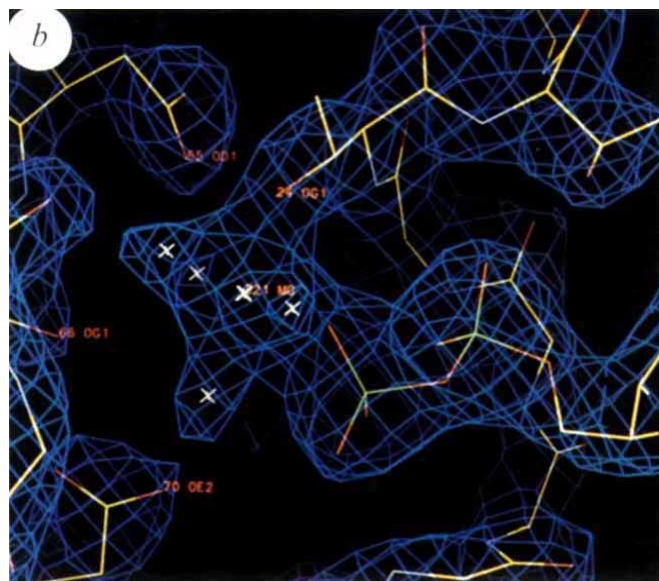


FIG. 3 The binding of $Mg \cdot GDP$ to Ran. *a*, Schematic drawing of the nucleotide-binding site with selected interactions (dashed lines) between Mg^{2+} , GDP and the protein with the corresponding distances (in Å). Additional interactions between main-chain and side-chain atoms stabilize the conformation of the residues shown, such as the interaction between T24 and D65. As expected, there are numerous electrostatic interactions between residues of the phosphate-binding loop (P-loop) (motif GxxxGKS/T; ¹⁷GDGGTGKT²⁴ in Ran) and the oxygens of the α - and β -phosphates that stabilize nucleotide binding. Mg^{2+} has the characteristic octahedral coordination pattern in which the hydroxyl group of T24 (S17 in Ras), one β -phosphate oxygen and two water molecules (W246 and W248), centred around the ion, form the square base of the octahedron and two further water molecules (W247 and W251) are positioned at its apices. As in Ras, D65 (D57 in Ras) interacts indirectly with Mg^{2+} via the first-coordination-sphere ligands water (W251) and T24. *b*, Part of the 2.3 Å electron density map contoured around 17.5% of the maximum, in the vicinity of the Mg^{2+} -binding site with the coordinating water molecules. Density for residues 65–70 (switch II in Ras) is well defined.



that changes its conformation when going from the GTP- to the GDP-bound state, our findings indicate that T66 and E70 might be important during the conformational change that is mediated by the loss of the γ -phosphate and the environmental change close to the Mg^{2+} ion. Whether the coordination of these residues to Mg^{2+} provides an explanation for the conservation of the DTAGQE motif in Ras-related proteins remains to be determined.

Another striking feature of Ran is the location of T42 (T35 in Ras). For Ras-related proteins, the involvement of this highly conserved residue in Mg^{2+} and γ -phosphate coordination has been considered to be central to the structural change occurring on GTP binding¹⁰ and confirmed by crystallographic analysis of other GTP-binding proteins^{13–16}. T42 is found at the C-terminal end of β_2E , 18 Å away from the Mg^{2+} -binding site. Thus assuming that the nature of the conformational change controlled by the loss of the γ -phosphate of GTP is conserved, one would have to postulate large movements of the region corresponding to the 'effector loop' in Ras to bring T42 close to the Mg^{2+} ion and γ -phosphate.

Regulatory proteins for Ran such as the guanine-nucleotide-exchange factor RCC1 (ref. 20) and ranGAP²¹, a Ran-specific GTPase-activating protein, have been described. Unlike with Ras, not much is known about the sites of their interactions with Ran, except that the mutation T24→N (analogous to S17→N in Ras) results in a dominant negative inhibitor which binds tightly to RCC1 in the presence of guanine nucleotides⁶, and that the ranGAP-stimulated GTPase is blocked by the mutation Q69→L (Q61→L in Ras^{22,23}). Possible effectors of Ran have also been described as RanBP (Ran-binding proteins)^{22,24} and it has been shown that the presence of GTP and an intact C terminus is necessary for this interaction²⁴. Considering the distance between the nucleotide and the C-terminal acidic tail, which we measured as 35 Å (β -phosphorus of GDP to the α -carbon of P210), it will

be interesting to see how the protein in the GTP-bound form accommodates RanBP. We are therefore focusing our attention on the structure on the GTP-bound form of Ran and complexes of Ran with regulatory proteins and effectors. □

Received 7 November 1994; accepted 31 January 1995.

1. Valencia, A., Chardin, P., Wittinghofer, A. & Sander, C. *Biochemistry* **30**, 4637–4648 (1991).
2. Bourne, H. R., Sanders, D. A. & McCormick, F. *Nature* **348**, 125–131 (1990).
3. Bischoff, F. R. & Ponstingl, H. *Proc. natn. Acad. Sci. U.S.A.* **88**, 10830–10834 (1991).
4. Moore, M. S. & Blobel, G. *Nature* **365**, 661–663 (1993).
5. Melchior, F., Paschal, B., Evans, J. & Gerace, L. *J. Cell Biol.* **123**, 1649–1659 (1993).
6. Kornbluth, S., Dasso, M. & Newport, J. *J. Cell Biol.* **125**, 705–719 (1994).
7. Ren, M., Coutavas, E., D'Eustachio, P. & Rush, M. G. *Molec. cell. Biol.* **14**, 4216–4224 (1994).
8. Clarke, P. R., Klebe, C., Wittinghofer, A. & Karsenti, E. *J. Cell Sci.* (in the press).
9. Milburn, M. V. et al. *Science* **247**, 939–945 (1990).
10. Wittinghofer, A. & Pai, E. *Trends biochem. Sci.* **16**, 383–387 (1991).
11. Pai, E. F. et al. *Nature* **341**, 209–214 (1989).
12. Jurnak, F. *Science* **230**, 32–36 (1985).
13. LaCour, T. F. M., Nyborg, J., Thirup, S. & Clark, B. F. C. *EMBO J.* **4**, 2385–2388 (1985).
14. Berchtold, H. et al. *Nature* **365**, 126–132 (1993).
15. Noel, J. P., Hamm, H. E. & Sigler, P. B. *Nature* **366**, 654–663 (1993).
16. Coleman, D. E. et al. *Science* **265**, 1405–1412 (1994).
17. Czworkoski, J., Wang, J., Steitz, T. A. & Moore, P. B. *EMBO J.* **13**, 3661–3668 (1994).
18. Evarsson, A. et al. *EMBO J.* **13**, 3669–3677 (1994).
19. Klebe, C., Nishimoto, T. & Wittinghofer, F. *Biochemistry* **32**, 11923–11928 (1993).
20. Ohtsubo, M., Okazaki, H. & Nishimoto, T. *J. Cell Biol.* **109**, 1389–1397 (1989).
21. Bischoff, F. R., Klebe, C., Kretschmer, J., Wittinghofer, A. & Ponstingl, H. *Proc. natn. Acad. Sci. U.S.A.* **91**, 2587–2591 (1994).
22. Coutavas, E., Ren, M., Oppenheim, J. D., D'Eustachio, P. & Rush, M. G. *Nature* **366**, 585–587 (1993).
23. Klebe, C., Bischoff, F. R., Ponstingl, H. & Wittinghofer, A. *Biochemistry* **34**, 639–647 (1995).
24. Lounsbury, K. M., Beddow, A. L. & Macara, I. G. *J. Biol. Chem.* **269**, 11285–11290 (1994).
25. Kabsch, W. *J. appl. Crystallogr.* **26**, 795–800 (1993).
26. Tong, L., de Vos, A. M., Milburn, M. V. & Kim, S.-H. *J. molec. Biol.* **217**, 503–516 (1991).
27. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. *Acta crystallogr. A* **47**, 110–119 (1991).
28. Brünger, A. T. *X-PLOR* (Version 3.1) (Yale Univ., New Haven, 1993).
29. Kabsch, W. & Sander, C. *Biopolymers* **22**, 2577–2637 (1983).
30. Kraulis, P. J. *J. appl. Crystallogr.* **24**, 946–950 (1991).

ACKNOWLEDGEMENTS. We thank I. Schlichting for reviewing the manuscript; R. Bischoff and N. Nassar for critical comments; F. Schmitz for help with the graphical representation; R. Bischoff, N. Nassar, T. Schweins, M. Geyer and A. Sanchez for discussion; P. Metcalf and G. Petsko for advice at the beginning of the project; and K. Holmes for encouragement. This work was supported by the National Institute of Cancer.

Structure and function of the multifunctional DNA-repair enzyme exonuclease III

Clifford D. Mol*, Che-Fu Kuo*†, Marla M. Thayer*, Richard P. Cunningham‡ & John A. Tainer*§

* Department of Molecular Biology, The Scripps Research Institute, 10666 North Torrey Pines Road, La Jolla, California 92037, USA

† Department of Biological Sciences, Center for Biochemistry and Biophysics, State University of New York-Albany, Albany, New York 12222, USA

THE repair of DNA requires the removal of abasic sites, which are constantly generated *in vivo* both spontaneously¹ and by enzymatic removal of uracil², and of bases damaged by active oxygen species, alkylating agents and ionizing radiation^{3,4}. The major apurinic/aprimidinic (AP) DNA-repair endonuclease in *Escherichia coli* is the multifunctional enzyme exonuclease III, which also exhibits 3'-repair diesterase, 3'→5' exonuclease, 3'-phosphomonoesterase and ribonuclease activities⁵. We report here the 1.7 Å resolution crystal structure of exonuclease III which reveals a 2-fold symmetric, four-layered $\alpha\beta$ fold with similarities to both deoxyribonuclease I⁶ and RNase H⁷. In the ternary complex determined at 2.6 Å resolution, Mn^{2+} and dCMP bind to exonuclease III at

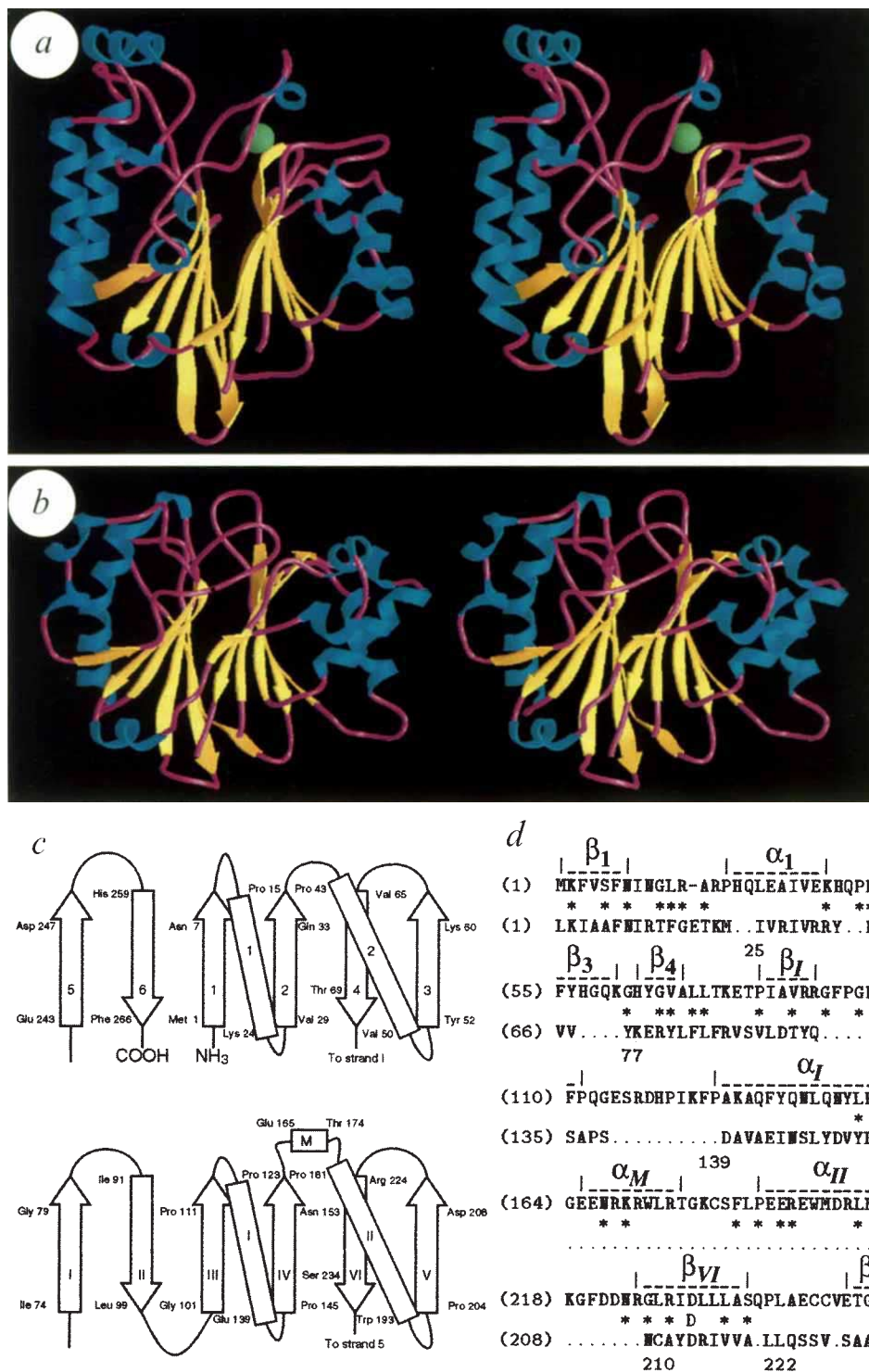
one end of the $\alpha\beta$ -sandwich, in a region dominated by positive electrostatic potential. Residues conserved among AP endonucleases from bacteria to man cluster within this active site and appear to participate in phosphate-bond cleavage at AP sites through a nucleophilic attack facilitated by a single bound metal ion.

We determined the exonuclease III structure by multiple isomorphous replacement (MIR) with three heavy-atom derivatives and refined it to 1.7 Å resolution with a crystallographic *R*-value of 0.190 (Table 1). The enzyme is a compact, globular $\alpha\beta$ protein of rough dimensions 55 Å × 50 Å × 45 Å consisting of two six-stranded β -pleated sheets flanked by four α -helices forming a four-layered $\alpha\beta$ -sandwich motif⁸ (Fig. 1a). This central four-layered $\alpha\beta$ -sandwich topology is surprisingly similar to the overall fold of the sequence-dependent endonuclease DNase I⁶ (Fig. 1b), with each β -sheet composed of two three-stranded anti-parallel motifs that meet to form two central parallel β -strands (Fig. 1c). The two β -sheets pack closely together with a mean

FIG. 1 Exonuclease III fold, topology and amino-acid sequence conservation. a, Exonuclease III stereo ribbon diagram with β -strands (yellow ribbons), with the arrow heads indicating the direction of the chain, α -helices (blue coils), loops (purple tubes) and the divalent metal ion (green sphere). The active site is located at the top of the figure at the base of the cleft between the two β -sheets. b, Deoxyribonuclease I stereo ribbon diagram with the same orientation and colouring scheme as for exonuclease III in a. The four-layered sandwich fold is apparent in both enzymes, but their respective DNA-binding surfaces, at the top of both structures, are clearly different. c, Exonuclease III polypeptide topology schematic. The β -strands are depicted as arrows with the

† Present address: Department of Structural Biology, Department 46Y, Building AP9A, Abbott Laboratories, One Abbott Park Road, Abbott Park, Illinois 60064, USA.

§ To whom correspondence should be addressed.



arrow head indicating the direction of the chain. α -Helices are shown as boxes. Each secondary structure element is numbered as it occurs in the amino-acid sequence with the β -strands and α -helices of the second β -sheet assigned Roman numerals, and the starting and ending residues of each as indicated. The DNA-binding surface is located at the top of each β -sheet. The 2-fold symmetry of the fold is apparent in the figure. To reconstruct the three-dimensional fold of the protein, rotate the top β -sheet by 180° around an axis perpendicular to the plane of the page and fold the page between the two β -sheets. *d*, Exonuclease III amino-acid sequence²¹, secondary structure, sequence conservation with the AP endonucleases and sequence alignment with regions of DNase I. The locations of the α -helices and β -strands

that make up the four-layered sandwich, as well as the protruding α -helix, α_M , are indicated. Also shown are those residues of DNase I¹⁰ that are structurally similar to exonuclease III. Starting residue numbers are indicated along the side for each row, and numbering along the bottom is with respect to the various non-contiguous segments of DNase I. Dashes indicate single amino-acid gaps, and dots in the DNase I sequence indicate the interconnecting loop regions that vary substantially from the AP endonucleases. Asterisks mark the amino-acid sequence identities between exonuclease III and the human²², murine²³ and bovine²⁴ AP endonucleases. The sequence alignment was generated using the MSA package²⁵ and optimized with respect to the known secondary structure of exonuclease III.

TABLE 1 Crystallographic data statistics

Parameters	Native	Mn-dCMP	Thimerosal	pHgAc*	KAu(CN) ₂
Structure determination					
Heavy atom (mM)	—	1	1	10	20
Soaking time (h)	—	36	13	8	43
Data collection device†	xen	ip	xen	xen	xen
Resolution (Å)	1.7	2.6	2.7	2.7	2.7
Data completeness (%)	92	95	99	98	91
<i>I</i> / σ at resolution limit	1.5	10.6	2.9	3.8	3.5
No. of observations	199,757	45,208	28,204	20,844	22,530
Unique reflections	28,817	9,468	7,830	7,766	7,248
<i>R</i> _{sym} (%)‡	8.5	6.7	10.0	8.3	7.2
<i>R</i> _{iso} (%)§	—	—	22.5	27.9	22.5
<i>R</i> _{Cullis}	—	—	0.50	0.50	0.53
Phasing power	—	—	2.00	1.99	1.81
No. of sites	—	1	2	3	1
Structure refinement					
Resolution Å	8.0–1.70	8.0–2.6			
Crystallographic <i>R</i> value¶ (reflections)					
All data	0.190 (26,638)	0.181 (8,166)			
<i>I</i> / σ > 3	0.179 (24,544)	0.178 (7,888)			
Total number of protein atoms	2,182	2,182			
Total number of DNA atoms	—	20			
Total number of metal atoms	—	1			
Total number of solvent molecules	257	182			
Average <i>B</i> value (Å ²)					
main chain	10.0	5.7			
side chain	14.2	8.8			
solvent	34.3	32.2			
dCMP	—	26.0			
Mn ²⁺	—	10.2			
R.m.s.d. bond lengths (Å)	0.012	0.016			
R.m.s.d. bond angles (°)	2.841	3.430			
R.m.s.d. dihedral angles (°)	24.97	25.62			
R.m.s.d. improper angles (°)	1.10	1.86			

Exonuclease III crystals are grown by vapour diffusion from solutions containing polyethylene glycol 4000 (Sigma) and high concentrations of imidazole-malate buffer, and belong to space group *P*3₁21, *a* = *b* = 107.8 Å, *c* = 42.8 Å (ref. 16). The structure was determined using the method of multiple isomorphous replacement, augmented by solvent levelling and, in the latter stages of the structure determination, by use of phase combination with the program σ_A ¹⁷. The amino-acid sequence was visible in and fit to both the 2.7 Å resolution MIR (figure-of-merit 0.62), and solvent-levelling MIR (figure-of-merit 0.85), electron density maps using the interactive graphics program XFIT¹⁸. Refinement was done using simulated annealing and positional refinement in X-PLOR¹⁹, and the correctness of the trace confirmed and any missing side chains built into the electron density of (2|*F*_o| – |*F*_c|), (|*F*_o| – |*F*_c|), and σ_A weighted |*F*_o| – |*F*_c| omit maps calculated with about 10% of the residues omitted from the phase calculation. Well ordered solvent molecules were either manually fitted, or automatically placed using a peak-picking program²⁰, and confirmed by molecular graphics. R.m.s.d., root-mean-square deviation.

* pHgAc, Phenylmercuric acetate.

† xen, Siemens multiwire area detector; ip-MAR Research Image Plate.

‡ *R*_{sym} is the unweighted *R* value on *I* between symmetry mates.

§ $R_{iso} = \sum (|F_{der}| - |F_{nat}|) / \sum |F_{nat}|$.

|| $R_{Cullis} = \sum [(|F_h| - (|F_{ph}| - |F_p|))] / \sum |F_h|$, for centric reflections.

¶ $R = \sum_{hkl} |F_{obs}(hkl) - F_{calc}(hkl)| / \sum_{hkl} F_{obs}(hkl)$.

distance of 6.0 Å between *Ca* atoms. The two distinct halves of the molecule, each containing topological units of secondary structure of the form $\beta\beta\beta\alpha\beta\alpha\beta\beta$, are related by an internal pseudo-2-fold rotation axis running along the sheet interface. Thus, the β -sheet β_1 – β_6 and α -helices α_1 – α_2 are pseudo-symmetric to β_1 – β_{VI} and α_1 – α_{II} (Fig. 1a, c). This pseudo-symmetry does not extend to the loops connecting the β -strands and α -helices, nor is it mirrored in the amino-acid sequence (Fig. 1d). Besides the α -helices within the four-layered sandwich, there is a three-turn α -helix (designated α_M) that protrudes from one end of the protein in the β_{IV} – α_{II} loop and is positioned for possible DNA major groove recognition (Fig. 1a, c, d). Whereas the enzyme's overall fold, which forms a long groove at one end of the $\alpha\beta$ -sandwich, is similar to DNase I, the loops surrounding this groove are quite different (Fig. 1a, b), and there is less than 20% overall sequence similarity between the two endonucleases (Fig. 1d).

The single divalent metal ion and nucleotide binding sites locate the exonuclease III active site at the bottom of the groove between the two β -sheets (Fig. 1a). The groove sides are formed by loops connecting β_5 to β_6 , β_1 to α_1 and β_3 to β_4 and the longer loops that connect β_{III} to α_1 , β_{IV} to α_{II} and β_V to β_{VI}

(Fig. 1c). Polar residues within the active site form hydrogen bonds with a number of ordered water molecules, particularly within a pocket surrounded by Gln 112, Asn 153, Tyr 109, Asn 7, His 259 and Trp 212 (Fig. 2a). In the ternary complex with Mn²⁺ and dCMP, the single Mn²⁺ ion is bound (~2.2 Å) to Glu 34 O ϵ 1 (Fig. 2b) at the bottom of this groove near the β -sheet interface. Other potential ligands (Asn 9 and Asn 7 O δ 1, Asp 258 O δ 2 and His 259 N ϵ 2) lack the correct orientations or bond lengths for metal ion ligation, but movements of Asp 258 could allow metal ion binding.

The possibility that exonuclease III may bind two metal ions, such as the Mg²⁺ and Zn²⁺ ions found in the DNA polymerase I 3',5'-exonuclease domain⁹, appears unlikely. No notable peaks are observed when native crystals are soaked in 10 mM Zn²⁺ solutions, and peaks are observed only at the same single site for soaks with Mn²⁺ or the calcium analogue Sm²⁺. In the ternary complex, the Mn²⁺ electron density (Fig. 2b) is twofold higher than that seen for the bound Mn²⁺ alone (data not shown), indicating that the metal ion binds with greater occupancy in the presence of nucleotide. The exonuclease III active site (Fig. 2a) is lined with charged and polar residues that interact with dCMP (Fig. 2b). Both Asn 153 and Gln 112 side chains hyd-

FIG. 2 Active-site electron density for exonuclease III and its ternary complex with Mn^{2+} and dCMP, and proposed reaction mechanism for cleavage of the P-O3' bond. *a*, Electron density for the active site of native exonuclease III. The σ_A -weighted map was calculated with coefficients $2F_o - F_c$, α_{calc} at a resolution of 1.7 Å, and is contoured at a level of 1.0σ . The position of the nucleophilic water molecule (HOH) was not included in the phase calculation. *b*, Electron density for the ternary complex of exonuclease III and Mn^{2+} and dCMP. Electron density is shown from a σ_A -weighted $F_{dCMP} - F_{metal-free}$ difference map at 2.0σ (blue contours) and at 4σ (pink contours). Native crystals were stabilized and washed before soaking in solutions containing 20 mM dCMP and 20 mM $MnSO_4$. The divalent metal-binding site also binds Mn^{2+} and Sm^{2+} in the absence of nucleotide (data not shown), but is occupied at a higher level when the nucleotide is present. The second highest feature in the difference maps is occupied by the dCMP phosphate group. *c*, Proposed reaction mechanism for the hydrolytic cleavage of the P-O3' bond. Asp 229 forms a hydrogen bond with His 259 and stabilizes the positive charge that develops when, acting as a general base, the histidine abstracts a proton from a water molecule, HOH 257, opposite the O3' atom of the scissile phosphate. The resultant nucleophilic hydroxide ion then attacks the phosphate group, resulting in an inversion of its configuration, as it proceeds through a penta-covalent transition state. The metal ion bound by Glu 34 interacts with the negatively charged phosphate group and aids the nucleophilic attack of the hydroxyl. This mechanism may apply to all three nucleolytic activities of exonuclease III: AP endonuclease, 3'→5'-exonuclease and RNase H activities.

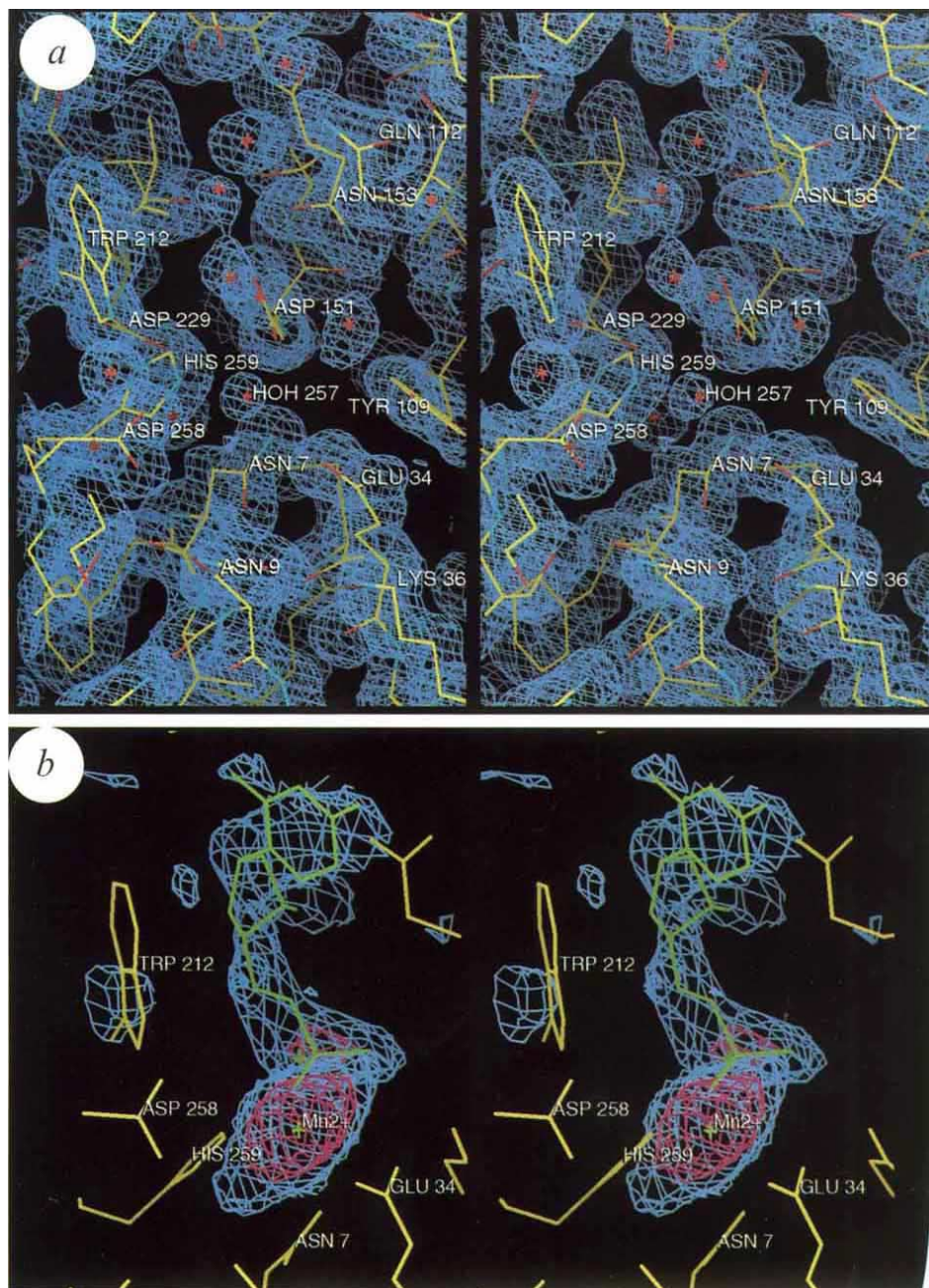
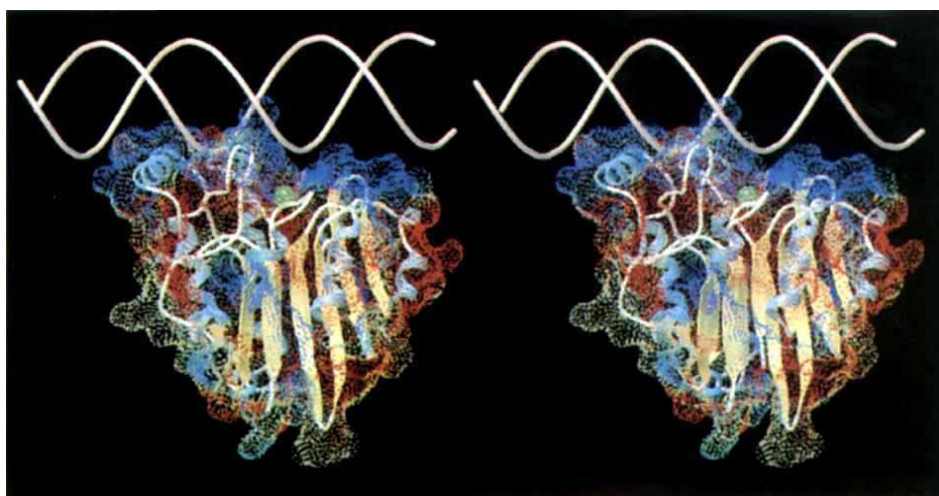


FIG. 3 Electrostatic surface and fold of exonuclease III. The electrostatic potential is mapped onto the solvent-accessible surface of the enzyme as colour-coded dots (blue, positive; red, negative; and neutral, green). Surface potential was calculated for the native metal-free structure with the program DELPHI (Biosym Technologies Inc.) using the protein partial charges, and a radius of 1.7 Å for the solvent, with the solute and solvent dielectric constants set to 2.0 and 80.0, respectively. The DNA backbone in the modelled exonuclease III–DNA complex (pink tubes) is shown with the exonuclease III ribbon diagram (yellow β -strands, blue α -helices and white loops) and single active-site metal ion (green sphere). The penetration by the protruding α -helix into the DNA major groove (upper left) and the interactions of the β_V – β_{VI} loop with the minor groove (top centre) are evident.



rogen bond to the nucleotide O3' position (distances of 2.8 and 3.3 Å, respectively), whereas the Asn 153 side chain also interacts with the 5'-phosphate group, which has four additional residue interactions (Asn 7, Asp 151, His 259 and Tyr 109). The Trp 212 indole ring stacks with the nucleotide sugar ring (Fig. 2b). These active site residues in exonuclease III are almost entirely conserved among the mammalian AP endonucleases (Fig. 1d).

These structural results suggest that the AP endonucleases act through a single metal-ion aided nucleophilic attack on the P–O3' bond (Fig. 2c). The proposed Asp-His-H₂O proton acceptors and donors are thus similar to the Asp-His-Ser catalytic triad observed in serine proteinase structures. The positively charged metal ion may function to orient the phosphate group, stabilize the transition state, and also polarize the P–O3' bond. The role of the Lewis acid, to protonate the O3' leaving group, may be played by the Asp 151 carboxylate (Fig. 2a) whose pK_a may be altered by its extensive hydrogen bonding network with Asn 7 and Asn 153. Interestingly, DNase I may have a similar endonuclease mechanism^{10,11}.

The exonuclease III electrostatic profile (Fig. 3) reveals that the metal ion binds in a small negatively charged pocket within an extended positively charged DNA-binding face that is in turn surrounded by negative potential reflecting the enzyme's net charge of –4 at neutral pH. This extensive electrostatically positive exonuclease III DNA-binding face (55 Å long) implies a larger interface for interaction with DNA than observed in the DNase I–DNA complexes^{10,11}. Yet, exonuclease III may mimic the DNA contacts made by the DNase I 'minor-groove binding loop': Arg 90 and Lys 36 can interact with the phosphodiester backbone, whereas Tyr 63 can stack with the sugar ring of the deoxyribose. However, two additional regions of exonuclease III may interact with duplex DNA. The protruding α -helix, α_M , is positioned for possible major-groove interactions, about one half turn of DNA (or six base pairs) from the active site, whereas the β_V – β_{VI} loop (Fig. 1c, d) could then interact with the minor groove. The entire enzyme would contact about 1,700 Å² of the DNA surface. However, some conformational change of the enzyme, the DNA or both would be needed for the DNA to fit both onto the protruding α -helix and into the dCMP site, and this requirement suggests a possible mechanism for the enzyme's specificity.

Recognition of an AP site might include interaction with an extra-helical base on the DNA strand opposite the AP site, perhaps in the pocket between Trp 212 and Gln 112, with the enzyme recognizing any unpaired base through Gln 112 or Asn 153 hydrogen bonding, base stacking with Trp 212, and differential replacement of bound water molecules (Fig. 2a, b). However, more specificity might result from AP site-promoted

distortion of the DNA sugar-phosphate backbone, as the enzyme appears designed to allow abasic sites to protrude into the active site for cleavage whereas the protruding α_M helix and surrounding loops keep the undamaged DNA backbone away from the catalytic site. Indeed, for RNase H activity, exonuclease III must be able to recognize distortions of the B-form double helix resulting from the RNA–DNA hybrid duplex. For exonuclease III acting as a 3'→5'-exonuclease, only a slight tilting of the DNA, with α_M acting as a pivot point, places the end of the DNA helix in the active site where the sugar–phosphate backbone could be cleaved.

Some structural and functional similarities shared by exonuclease III and DNase I also extend to RNase H, which has an identical enzymatic activity with exonuclease III and a similar fold around the metal-binding site¹². The RNase H β -sheet superimposes onto β -strands IV, III, II and VI of exonuclease III and results in superimposed active sites that share putative catalytic residues: RNase H Asp 70 and His 124 versus exonuclease III Asp 229 and His 259. Of the five histidines in *E. coli* RNase H, mutagenesis of His 124 to Ala reduces k_{cat} to 3.3% of the wild-type enzyme¹³, and a rigorous examination of metal binding to *E. coli* RNase H¹⁴ suggests one metal ion is bound at the active site, as we also find for exonuclease III. Thus, the RNases H may act through a hydrolytic, single metal ion mechanism similar to the Asp-His-H₂O mechanism we propose for exonuclease III. Overall, these structural and functional similarities of exonuclease III with both DNase I and RNase H apparently link these enzymes into a single structural family of hydrolytic nucleases.

The refined, high-resolution structure of the first DNA repair AP endonuclease reveals a new use for the $\alpha\beta$ -sandwich fold in DNA base-excision repair. The enzyme's Mn²⁺–dCMP complex suggests that exonuclease III and the homologous mammalian AP endonucleases cleave the P–O3' bond at AP sites through a nucleophilic attack aided by a single divalent metal ion. The adjacent locations of a protruding basic α -helix for DNA major-groove interactions and of a protruding loop for minor-groove binding might aid specificity for distorted duplex DNA. Structural similarities of exonuclease III with both DNase I and RNase H at and around their active sites suggest that all three nucleic acid hydrolytic enzymes have converged upon a functionally appropriate protein fold, encompassing a single divalent metal ion active site allowing similar hydrolytic mechanisms, whereas distinct surrounding loop regions determine substrate binding and catalytic specificity. A similar $\alpha\beta$ -sandwich scaffold positions structural elements including a protruding basic helix for DNA interaction in the p53 tumour suppressor–DNA complex¹⁵. The occurrence of the $\alpha\beta$ -sandwich structure in a

DNA hydrolytic enzyme, a transcription activator protein and also in a DNA repair enzyme indicates that this motif may be a recurrent structural theme in the molecular biology of proteins that interact with DNA. □

Received 9 November 1994; accepted 26 January 1995.

1. Lindahl, T. *Nature* **362**, 709–715 (1993).
2. Taylor, A. F. & Weiss, B. J. *Bact.* **151**, 351–357 (1982).
3. Sakumi, K. & Sekiguchi, M. *Mutat. Res.* **236**, 161–172 (1990).
4. Doetsch, P. W. & Cunningham, R. P. *Mutat. Res.* **236**, 173–201 (1990).
5. Dimple, B. & Harrison, L. A. *Rev. Biochem.* **63**, 915–948 (1994).
6. Oefner, C. & Suck, D. J. *molec. Biol.* **192**, 605–632 (1986).
7. Katayanagi, K. et al. *Nature* **347**, 306–309 (1990).
8. Richardson, J. S. *Adv. Prot. Chem.* **34**, 167–339 (1981).
9. Beese, L. S. & Steitz, T. A. *EMBO J.* **10**, 25–33 (1991).
10. Lahm, A. & Suck, D. J. *molec. Biol.* **221**, 645–667 (1991).
11. Weston, S. A., Lahm, A. & Suck, D. J. *molec. Biol.* **226**, 1237–1256 (1992).
12. Katayanagi, K. et al. *Nature* **347**, 306–309 (1990).
13. Kayana, S. et al. *Eur. J. Biochem.* **198**, 437–440 (1991).
14. Katayanagi, K., Okumura, M. & Morikawa, K. *Prot. Struct. Funct. Genet.* **17**, 337–346 (1993).
15. Cho, Y., Gorina, S., Jeffrey, P. D. & Pavletich, N. P. *Science* **265**, 346–355 (1994).
16. Kuo, C.-F., McRee, D. E., Cunningham, R. P. & Tainer, J. A. *J. molec. Biol.* **229**, 239–242 (1993).
17. Read, R. J. *Acta crystallogr.* **A42**, 140–149 (1986).
18. McRee, D. E. *J. molec. Graphics* **10**, 44–46 (1993).
19. Brünger, A., Kuriyan, J. & Karplus, M. *Science* **235**, 458–460 (1987).
20. Borgstahl, G. E. O., Rogers, P. H. & Arnone, A. J. *molec. Biol.* **236**, 817–830 (1994).
21. Saporito, S. M., Smith-White, B. J. & Cunningham, R. P. *J. Bact.* **170**, 4542–4547 (1988).
22. Dimple, B., Herman, T. & Chen, D. S. *Proc. natn. Acad. Sci. U.S.A.* **88**, 11450–11454 (1991).
23. Seki, S. et al. *J. biol. Chem.* **266**, 20797–20802 (1991).
24. Robson, C. N., Milne, A. M., Pappin, D. J. C. & Hickson, I. D. *Nucleic Acids Res.* **19**, 1087–1092 (1991).
25. Lipman, D. J., Altschul, S. F. & Kececioglu, J. D. *Proc. natn. Acad. Sci. U.S.A.* **86**, 4412–4415 (1989).

ACKNOWLEDGEMENTS. We thank H. E. Parge, E. D. Getzoff, T. J. Macke, M. E. Pique, D. E. McRee, M. A. McTigue, G. E. O. Borgstahl, B. R. Crane, A. S. Arvai and R. Bishop for helpful discussions. This work was supported by the NIH (R.P.C. and J.A.T.) and by the National Science Foundation (R.P.C.).

Separate domains of p21 involved in the inhibition of Cdk kinase and PCNA

Junjie Chen*, Peter K. Jackson†, Marc W. Kirschner† & Anindya Dutta*‡

* Department of Pathology, Division of Molecular Oncology, Brigham and Women's Hospital, Harvard Medical School, 75 Francis Street, Boston, Massachusetts 02115, USA

† Department of Cell Biology, Harvard Medical School, 25 Shattuck Street, Boston, Massachusetts 02115, USA

The protein p21 (WAF1, CIP1 or sdi1), induced by the tumour-suppressor protein p53, interacts with and inhibits two different targets essential for cell-cycle progression^{1–8}. One of these is the cyclin-Cdk family of kinases and the other is the essential DNA replication factor, proliferating-cell nuclear antigen (PCNA). We report here that separate domains of p21 are responsible for interacting with and inhibiting the two targets. An amino-terminal domain inhibits cyclin-Cdk kinases and a carboxy-terminal domain inhibits PCNA. Using these separated domains, we have determined that p21 inhibits different biological systems through different targets. The PCNA-binding domain is sufficient for inhibition of DNA replication based on simian virus 40, whereas the Cdk2-binding domain is sufficient for inhibition of DNA replication based on *Xenopus* egg extract and for growth suppression in transformed human cells.

We have successfully separated the domains of p21 responsible for inhibiting Cdk2 and PCNA (Fig. 1). When the N- or C-terminal halves of p21 (p21N or p21C) were expressed separately

as GST fusion proteins, p21N alone physically associated with bacterially expressed Cdk2 kinase polypeptide in a standard 'pull-down' assay. Further, p21 and p21N, but not p21C, inhibited the kinase activity of cyclin A-Cdk2.

In contrast, the ability to interact with PCNA is the function of the C-terminal half of p21 (p21C). In a pull-down assay, bacterially expressed PCNA was quantitatively bound by p21 and p21C (Fig. 2). As reported, GST-p21 inhibited the simian virus 40 (SV40)-based *in vitro* DNA replication reaction (Fig. 2), and this inhibition was reversed by the addition of PCNA or a fraction of cell extract containing PCNA (refs 7, 8 and data not shown). The ability to inhibit the SV40-based DNA replication reaction required the PCNA-binding domain present

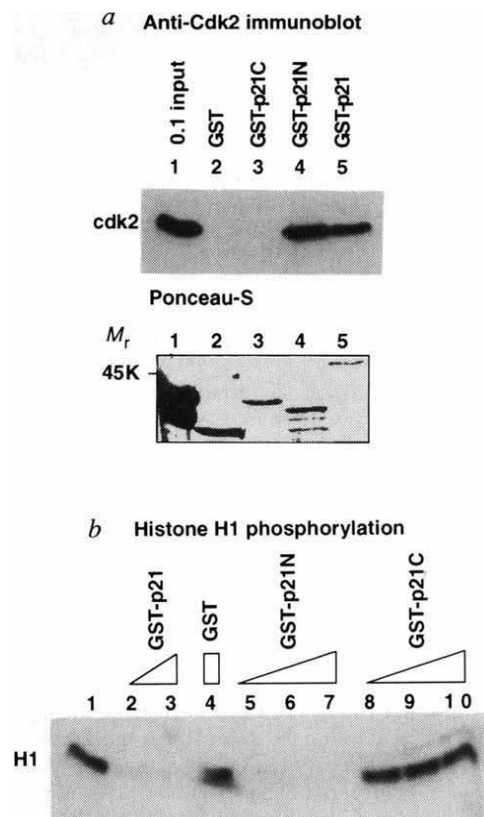


FIG. 1 The N-terminal half of p21 binds to and inhibits Cdk2 kinase. **a**, Immunoblot probed with anti-Cdk2 antibody shows which GST fusion proteins bind Cdk2. Ponceau S stain of the same blot demonstrates the relative amounts of the various GST fusion proteins present in each lane. The GST lane also contains M_r markers, and the 45K marker is indicated. The intense band in the 0.1 input lane is the carrier protein casein. **b**, Autoradiogram of proteins showing the phosphorylation of histone H1 by cyclin A-Cdk2 kinase with the following additions. Lane 1, no addition; 2, 3, 20 ng and 100 ng GST-p21; 4, 100 ng GST; 5–7, 20 ng, 100 ng, 200 ng GST-p21N; 8–10, 20 ng, 100 ng, 200 ng GST-p21C.

METHODS. Plasmids expressing GST fusion proteins in *Escherichia coli* were made by PCR of p21 and its deletion derivatives as *Bam*HI–*Sall* fragments, and cloning into the pGEX-5X-3 (Pharmacia). GST-p21 contains the whole coding region of human p21 gene (amino acids 1–164), GST-p21N contains amino acids 1–90, and GST-p21C contains amino acids 87–164. All p21 coding regions were preceded by a methionine. The expression and binding of GST fusion proteins to glutathione-agarose beads have been described before¹⁶. Cdk2 protein was expressed from *E. coli* using pET-Cdk2. The pull-down assay was done as previously described¹⁶, except that washing was done with buffer A. Histone H1 kinase assays were for 20 min at 30 °C using 10 ng of insect cell-expressed GST-cyclin A-Cdk2 complex and 8 µg of histone H1 in 20 µl kinase buffer¹¹. The phosphorylation of histone H1 was visualized by electrophoresis and autoradiography.

‡ To whom correspondence should be addressed.

in p21C (Fig. 2). Although the concentration of p21N added was sufficient to inhibit completely the endogenous histone H1 kinase activity in the extracts (data not shown), SV40 DNA replication was not inhibited.

p21 also inhibits DNA replication of sperm in interphase extracts derived from *Xenopus* eggs (P.K.J., S. Chevalier, M. Phillippe and M.W.K., manuscript submitted; also ref. 9). p21N inhibits DNA replication in the *Xenopus* extract at the same concentrations as p21 (Fig. 3a), which are similar to those of

cyclin E-Cdk2 (100 nM). Thus, in *Xenopus* extracts it appears that cyclin-Cdk2 rather than PCNA is limiting and is inhibited by p21. High concentrations of p21C, approximating that of PCNA in the extract (10 μ M), inhibited *Xenopus* DNA replication. Therefore, PCNA is required for double-stranded DNA replication and can be inhibited by p21, but is not the limiting factor inhibited by the addition of p21. In contrast to double-stranded DNA replication, DNA synthesis on single-stranded DNA was not inhibited by the Cdk inhibitory p21N domain but was inhibited by p21C at concentrations approaching that of PCNA (Fig. 3b). Therefore, active Cdk kinase is required specifically for DNA synthesis on double-stranded DNA, whereas PCNA is required for DNA synthesis on both types of substrates.

The question arises as to which of the two targets, Cdks or PCNA, mediates the growth suppression function of p21 in p53-null transformed cells in culture (Fig. 4a). As reported by others,

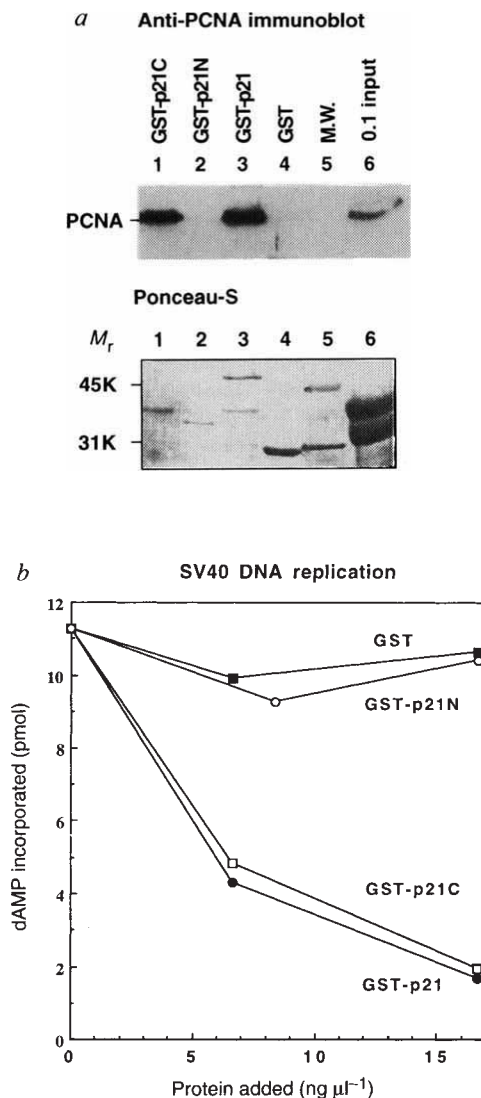


FIG. 2 The C-terminal half of p21 binds to PCNA and inhibits SV40-based DNA replication reaction. **a**, Immunoblot probed with anti-PCNA antibody (Oncogene Science) shows which GST fusion proteins bind PCNA. Ponceau S stain of the same blot demonstrates the relative amounts of the various GST fusion proteins present in each lane. **b**, Titration of various fusion proteins into SV40 DNA replication reactions containing SV40 T-antigen and a crude cell extract from human 293 cells. DNA synthesis is expressed as pmol of dAMP incorporated in a 50 μ l reaction in 1 h.

METHODS. PCNA protein was expressed from *E. coli* using pET-PCNA. Pull-down assay was exactly as described in Fig. 1. SV40 DNA replication was as previously described¹¹. pSV011 (180 ng) was replicated in a 30 μ l reaction containing 100 ng SV40 T-antigen (obtained from baculovirus expression system and already phosphorylated on the Cdk substrate site), 50 μ g of S100 extract from asynchronous 293 cells, and other indicated components. Crude extracts and T-antigen were preincubated on ice for 30 min with various GST fusion proteins without plasmid DNA, then replication reactions were done by mixing with plasmid DNA and incubation at 37 °C for 1 h.

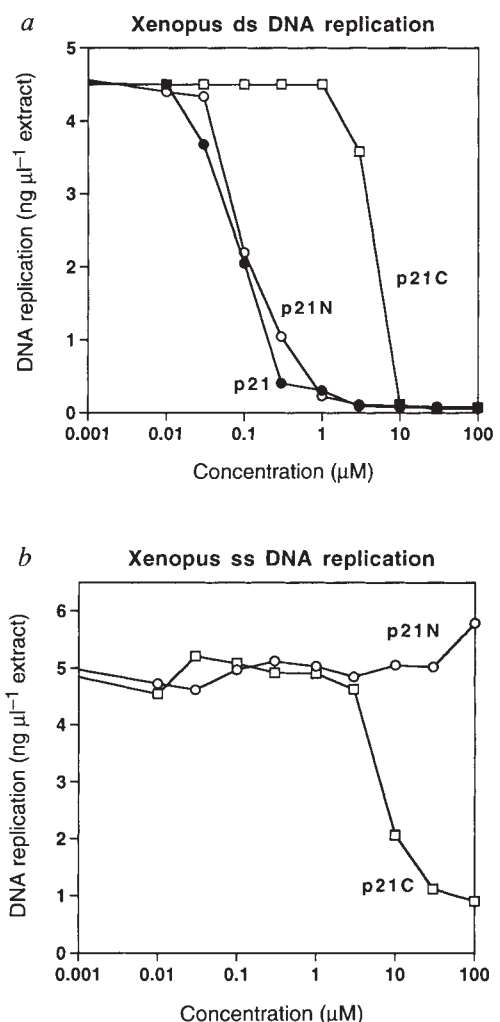


FIG. 3 **a**, p21N inhibits double-stranded sperm DNA replication in *Xenopus* egg extracts at similar concentrations as full-length p21. p21C inhibits DNA replication only when added at concentrations approaching that of PCNA in the *Xenopus* extracts. **b**, p21C but not p21N inhibits DNA synthesis on single-stranded DNA at concentrations similar to that required to inhibit DNA replication on double-stranded DNA.

METHODS. Interphase extracts were prepared from activated *Xenopus* eggs as described (P.K.J. et al., manuscript submitted). Briefly, 10 μ l reactions containing 5 ng μ l⁻¹ sperm DNA (or single-stranded M13 DNA) and [α -³²P]dATP were mixed with varying concentrations of bacterially expressed GST-fusion proteins, incubated for 2.5 h at 23 °C and assayed for TCA-precipitable counts. Amount of DNA replicated was calculated using the known specific activity of dATP in the reaction.

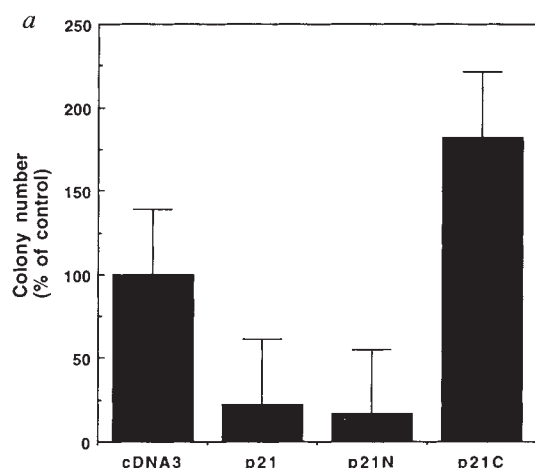
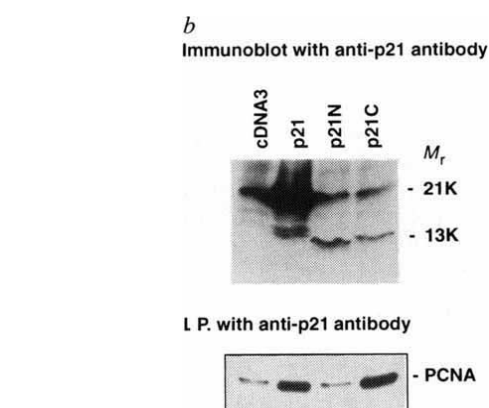


FIG. 4 a, p21N inhibits cell growth whereas p21C does not. Number of G418-resistant colonies obtained after transfection of SaOs2 cells with equal amounts of indicated plasmids is shown relative to the number obtained with vector control alone (cDNA3). Mean and standard deviation of five independent transfections. b, Upper, immunoblot probed with anti-p21 polyclonal antibody shows p21 and its derivatives were expressed *in vivo*. The 21K polypeptide seen in all lanes could be monkey p21. Exogenous proteins of 21, 13 and 13K are also seen in cells transfected with p21, p21N and p21C, respectively. Lower, the C-terminal half of p21 formed a complex with PCNA *in vivo*. Transiently transfected cell lysates were immunoprecipitated with anti-p21 polyclonal antibody and the precipitate immunoblotted with anti-PCNA antibody. All lanes contain PCNA presumably co-immunoprecipitated with monkey p21. Cells expressing p21 and p21C co-immunoprecipitate larger quantities of PCNA because of

plasmids expressing p21 established fewer colonies compared to control vector plasmids³. The N-terminal cyclin-Cdk2 inhibitory domain (p21N) also suppressed growth, and the C-terminal PCNA inhibitory domain (p21C) did not, suggesting that cyclin-Cdks are the primary targets of p21 in transfected cells. An immunoblot of extracts of transfected COS cells shows that human p21, p21N and p21C are expressed from these plasmids and that the exogenous human p21 and p21C proteins associated with and co-immunoprecipitated cellular PCNA (Fig. 4b). The PCNA-interacting portion of p21 (p21C) reproducibly stimulated the number of colonies obtained upon transfection, but the significance of this is unclear.

Having shown that the inhibition of cyclin-Cdk and inhibition of PCNA are independently executed by two different domains of p21, we could now determine which target mediates the effect of p21 in each biological system. Cdk kinases are the primary target for inhibition of double-stranded DNA replication in *Xenopus* extracts by p21 and for growth suppression in transfected cells. In contrast, PCNA is the limiting target in the SV40 replication reaction. The differences between the various systems suggest potential Cdk substrates for activation of DNA replication at the G1-S transition. The insensitivity of SV40 replication of p21N is paradoxical in view of the stimulation of replication in G1 extracts by Cdk kinase^{10,11}. The reported stimulation by Cdk kinase was not mediated through the phosphorylation of T antigen, because those experiments (like the



association of PCNA with the exogenous protein. METHODS. a, SaOs2 cells (p53 null human osteosarcoma cells) were transfected with 5 µg of the indicated plasmid and 10 µg of salmon sperm carrier DNA by the calcium-phosphate method¹¹. After 14–21 days of selection with G418 at 400 µg ml⁻¹ the number of colonies per transfection were counted. cDNA3 (Invitrogen) is the control vector which expresses genes inserted downstream from a cytomegalovirus promoter and contains a neomycin phosphotransferase gene. p21, p21N and p21C plasmids were made by transferring *Bam*HI-*Xho*I fragments from the corresponding pGEX 5X-3-derived plasmids (Fig. 1) into the same sites of cDNA3. b, COS cells (T-antigen transformed monkey kidney epithelial cells) were transfected as above with the indicated plasmids and collected 60 h later. Cell lysis and immunoprecipitation were done as previously described¹⁶. For the upper and lower panels, respectively, 110 and 1,100 µg of each cell lysate was used.

ones reported here) used T antigen which was already phosphorylated on the Cdk substrate site, threonine 124. One explanation could be that Cdk kinase is required specifically to overcome G1-specific factors and hence is dispensable for replication in extracts from asynchronous cells. In *Xenopus* extracts, potential targets for Cdk2 could be the hypothetical G1-specific factors, chromatin or origin-binding proteins (the T-antigen equivalent). Also, the failure of p21N to inhibit DNA synthesis on single-stranded DNA in *Xenopus* extracts suggests that Cdks are primarily required for the unwinding of and initiation of DNA synthesis on double-stranded DNA.

Inhibition of PCNA by p21 (and p21C) is not apparently required in the particular growth suppression assay used in this report^{6,12,13}. Transformed cells may have an excess of PCNA, so that higher levels of p21C are required to inhibit PCNA function. Inhibition of PCNA may also be more important for the p21-mediated temporary switch from DNA replication to repair following genotoxic damage¹². The ability of p21N alone to inhibit cell-cycle progression has interesting implications for the activity of other Cdk-inhibitory and cell-cycle blocking polypeptides like p27 (upregulated by transforming growth factor-β (TGF-β)) which show homology with p21 in the region contained in p21N^{14,15}. If the Cdk2-interacting domain of p21 can be narrowed down further, it may be possible to design small molecules that mimic the structure of this peptide. Such chemicals could suppress growth of cancers with loss of functional p53 or potentiate the growth-suppressive effect of TGF-β. □

Received 19 December 1994; accepted 27 January 1995.

- Xiong, Y. et al. *Nature* **366**, 701–704 (1993).
- Harper, J. W., Adami, G. R., Wei, N., Keyomarsi, K. & Elledge, S. J. *Cell* **75**, 805–816 (1993).
- el-Deiry, W. et al. *Cell* **75**, 817–825 (1993).
- Gu, Y., Turck, C. W. & Morgan, D. O. *Nature* **366**, 707–710 (1993).
- Noda, A., Ning, Y., Venable, S. F., Pereira, S. O. & Smith, J. R. *Exp. Cell Res.* **211**, 90–98 (1994).
- el-Deiry, W. et al. *Cancer Res.* **54**, 1169–1174 (1994).
- Waga, S., Hannon, G. J., Beach, D. & Stillman, B. *Nature* **369**, 574–578 (1994).
- Flores-Rozas, H. et al. *Proc. natn. Acad. Sci. U.S.A.* **91**, 8655–8659 (1994).
- Strausfeld, U. P. et al. *Curr. Biol.* **4**, 876–883 (1994).
- D'Urso, G., Marraccino, R. L., Marshak, D. R. & Roberts, J. M. *Science* **250**, 786–791 (1990).

- Dutta, A. & Stillman, B. *EMBO J.* **11**, 2189–2199 (1992).
- Li, R., Waga, S., Hannon, G. J., Beach, D. & Stillman, B. *Nature* **371**, 534–537 (1994).
- Dulic, V. et al. *Cell* **76**, 1013–1023 (1994).
- Polyak, K. et al. *Cell* **78**, 59–66 (1994).
- Toyoshima, H. & Hunter, T. *Cell* **78**, 67–74 (1994).
- Dutta, A., Ruppert, J. M., Aster, J. C. & Winchester, E. *Nature* **365**, 79–82 (1993).

ACKNOWLEDGEMENTS. We thank J. W. Harper, S. Elledge, D. Beach, B. Stillman, H. Piwnicka-Worms and M. Solomon for various reagents and E. Winchester and J. Morrow for technical support. This work was supported by grants and fellowships from the NIH, American Cancer Society, the Massachusetts Division of the American Cancer Society, US Armed Forces Medical Research Command and the Life Sciences Research Foundation.

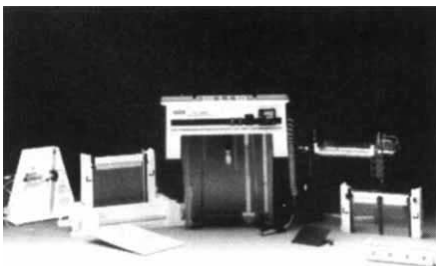
More than CAM and clathrin

Cells and tissues are under scrutiny for our cell biology feature. This weeks episode has five new protein kinases, mammalian expression vectors, cell death detection ELISA and proliferating human mesangial cells.

KAYIMA Biomedical has a new product line for **free-radical research** — carazostatin, pyrrolostatin and NCO-700 should have applications in the research of oxidative damage to biological tissues by free-radicals (*Reader Service No. 100*). Carazostatin and pyrrolostatin are strong inhibitors of lipid peroxidation induced by free-radicals. Carazostatin shows *ex vivo* inhibition of lipid peroxidation in mouse plasma and pyrrolostatin shows protective antihypoxic activity in mice. NCO-700 is a thiol protease inhibitor that inhibits oxidant production by fMLP-stimulated PMNs and scavenges reactive oxygen.

Clonetics has released the RenalPack-MS a cell system containing **cryopreserved or proliferating normal human mesangial cells**, optimized mesangial cell growth medium and subculture reagents for the propagation and subculture of the cells *in vitro* (*Reader Service No. 101*). The system can be used for immunohistochemical and biochemical studies, the study of filtration residues, peptide cytokines and growth factors, inflammatory response to glomerular injury and mesangial cell proliferation and differentiation.

The Bio-Rad D Gene system is designed for simple and efficient **mutation detection** (*Reader Service No. 102*). This is a denaturing gradient gel electrophoresis (DGGE) technique designed to identify single base mutations. The manufacture claims that it can provide up to 99 per cent efficiency. The system utilizes a vertical electrophoresis cell that separate wild-type and mutant fragments.



D Gene — mutation detection system.

The process uses a polyacrylamide gel with a denaturing gradient perpendicular to the electric field at a controlled temperature. There are two gel sizes available, a model 475 gradient delivery system, new MacMelt software (for predicting melt patterns), a reagent control kit and a reagent electrophoresis kit.

For the **identification of proliferation-associated antigens**, Coulter has released Clone p105 and Clone p120 (*Reader Service No. 103*). These products are intended to detect proliferation antigens that are upregulated during cell proliferation including oncoproteins and nuclear-associated antigens. Proliferative activity may have an important prognostic significance in several types of cancer, such as non-Hodgkin's lymphoma, neuroblastoma and cancers of the breast, colon, lung, ovary and bladder.

■ Protein products

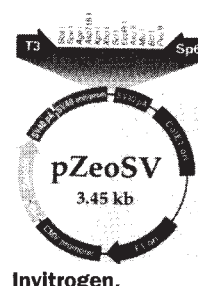
New England Biolabs has expanded its line of protein tools with the introduction of five new **protein kinases**: casein kinase II, cyclic AMP-dependent protein kinase, glycogen synthase kinase 3 and p34^{cdc2}/cyclin B (*Reader Service No. 104*). These proteins are expressed from recombinant sources, and are purified close to homogeneity and tested for contaminating activity. Related new products include protein phosphatase 1 and its specific inhibitor, I-2.

Among the latest products to be introduced to the Santa Cruz Biotechnology range of **signal transduction reagents** are members of the Id family of helix-loop-helix proteins (*Reader Service No. 105*). These proteins appear to regulate negatively the DNA binding of bHLH proteins; expression of each of the Id proteins is strongly dependent on growth factor activation. The new proteins are affinity-purified rabbit polyclonal Id1 (C-20), Id2 (C-20), Id3 (C-20) and Id4 (L-20). Also offered are new members of the Stat family (signal transducers and activators of transcription) — Stat4 (L-18) and Stat4 (C-20), which has been suggested to form homodimers as well as heterodimers with the other members of the Stat family.

Oxford Glycosystems now offers **phosphatidyl inositol specific phospholipase C (PI-PLC)** as a recombinant enzyme isolated from *Bacillus subtilis* transfected with the PI-PLC gene from *Bacillus thuringiensis* (*Reader Service No. 106*). This recombinant enzyme can be produced in bulk quantities and is intended for detection of certain GPI membrane anchors, structure/function studies, pharmaceutical research and therapeutic and vaccine development. The product is supplied in a neutral buffer and is protease free. A number of free pack sizes are available.

■ Gene expression

Invitrogen has introduced pZeoSV, the newest member in a line of **mammalian expression vectors** (*Reader Service No. 107*). This vector carries the Zeocin resistance gene for selection in both *Escherichia coli* and mammalian cell lines. By having only one selectable marker, the overall vector size of pZeoSV is only 3.45 kb, allowing more efficient ligation, transformation and transfection. pZeoSV contains the strong promoter enhancer from the early region of SV40 and is said to allow high-level expression of the desired gene in a variety of mammalian cell lines. The vector also carries a multi-ple cloning site for simplified cloning, the SV40 polyadenylation signal for efficient transcription and an f1 origin of replication for rescue of single-stranded DNA.



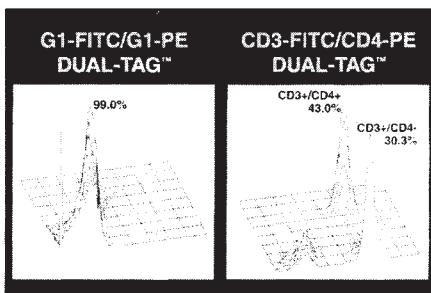
New antisense technologies are designed to permit control of cellular behaviour by manipulation of the expression of specific genes (*Reader Service No. 108*). TCS Biologicals has now introduced a stable **antisense oligonucleotide** preparation called Antisense retinoblastoma (Rb) which is directed against the retinoblastoma gene. Expression of the Rb gene product is a key control element which inhibits eukaryotic cells from entering mitosis. The product can simply be added to the medium to down regulate Rb gene expression and manufacturer states that the cell proliferation is then stimulated by three to five times. The effects of the antisense nucleotide is said to be reversible and cells revert to their normal doubling time when it is removed. It can be used to promote expansion of primary cultures and 'difficult' cell types including embryonic stem cells.

■ Cell cycle research

Promega's line of products for cell cycle research has the tools necessary to investigate cell-cycle control or identify potential anti-mitotic and anti-cancer drugs (*Reader Service No. 109*). The new product line includes pp34^{cdc2} kinase (eukaryotic cell-cycle study), olomoucine (pp34^{cdc2}/cyclin B complex inhibitor), cdc2 protein kinase peptide substrate (a synthetic peptide, specific for *in vitro* phosphorylation sites of histone H1) and Anti-pp34^{cdc2} kinase N-terminus (a polyclonal antibody for the active dephosphorylated kinase).

phorylated form of the protein). Also offered is p9^{Cks1H2} agarose that binds cell cycle-dependent protein kinases such as pp34^{cdc2} kinase, and to affinity purify this protein from human tissue and yeast extracts.

Boehringer Mannheim's new **cell death detection ELISA kit** is designed for determination of apoptosis — physiologically determined or 'programmed' cell death (Reader Service No. 110). The kit measures the mononucleosomes and oligonucleosomes released into the cytoplasm of cell populations during naturally occurring or experimentally generated apoptosis. Results are said to be obtained in 5–6 h, and



Dual-Tag cell death detection ELISA kit.

the assay does not generate radioactive waste. The kit's sandwich ELISA detects only the histone-associated DNA (nucleosomes, for example) that accompanies apoptosis, not free DNA.

■ Imaging and assays

Perseptive Biosystems has developed CytoFluor II for **quantitative fluorescent assays** (Reader Service No. 111). This instrument can help investigate tissue interactions, including inflammation and tumour formation, by measuring cell adhesion. The manufacturer states that this is done without altering the interaction between CAMs, their receptors, blocking agents, or analogues. In this assay, the fluorescent signal is emitted from cells that have had the volume of the cytosol labelled. Fluorescently labelled cells are added to a 96-well plate coated with monolayers of unlabelled cells such as endothelial or epithelial cells, or extracellular matrix proteins, known to have CAM-like properties. Labelled cells are allowed to adhere and unbound cells are washed away. The fluorescence is measured to determine the number of remaining adherent cells.

Life Science Resources has announced InSitu, an **image acquisition and analysis system** for molecular and cellular biology laboratories using immunocytochemistry, autoradiography and *in situ* hybridization (Reader Service No. 112). This system is designed for optical detection technologies capable of reporting quantity and localiza-

tion of proteins, antigens, surface receptors and genetic material. It provides quantitative densitometry on fixed cells and tissues, slices, sections and gels. InSitu is suitable for the detection of radioisotopic markers detected by autoradiography as well as light emitting optical reporters coupled to fluorescent, bio- or chemiluminescent molecular probes. A variety of imaging detector resolutions can be utilized and true colour, 24-bit image acquisition is also offered.

Sigma Immunochemicals has introduced seven new Dual-Tag **double-labelled monoclonal antibodies to human leukocyte markers** for flow cytometry (Reader Service No. 113). The CD Dual-Tag reagents available are human CD3-FITC/CD4-PE, human CD3-FITC/CD8-PE, human CD3-FITC/CD19-PE, human CD3-FITC/CD57-PE, human CD4-FITC/CD8-PE, human CD5-FITC/CD19-PE and human CD45-FITC/CD14-PE. The mixed reagents of FITC/PE labelled CD antibodies provide fluorescent labelling of cell-surface antigens and permit subsets of leukocyte populations to be analysed and sorted. Dual-Tag reagents are designed for characterization of subtypes of T-cell leukaemias and lymphomas; HIV infections, AIDS, and other associated diseases; identification and enumeration of helper/inducer, helper/suppressor T-cell subclasses and transplant studies.

■ Cell culture and refinement

Specialized Culture Technologies has introduced the POC-Culture chamber — it is designed to allow the **culture of many cell types** under stationary or continuous perfusion conditions on glass or permeable, hydrophilic or hydrophobic Teflon foil growth surfaces (Reader Service No. 114). In addition, the chamber can be placed in an incubator or directly on a heated microscope stage. The chamber has the same dimensions as a Terasaki plate and an observation area of about 6.6 cm². It has been designed to accommodate the new high-resolution oil-immersion objective, so that the refractive index of the growth surface does not present a problem.

The Harvest Mouse bioreactor from Serotec is designed for **monoclonal antibody production**, making affordable, hollow-fibre technology readily available (Reader Service No. 115). This reactor is said to require only a basic level of experience in tissue culture techniques for effective operation. The process is designed to be simple, self-contained and to run in a standard CO₂ incubator. The bioreactor can be autoclaved. □

These notes are compiled by Brendan Horton from information provided by the manufacturers. For more details, fill in the reader service card bound inside the journal.

ADVERTISEMENTS

Custom PEPTIDE Synthesis in sets of 95 + 1 control peptide

for screening of B and T-cell epitopes,
receptor binding studies etc.

Less than \$70/peptide!
(up to 18-mer, 2 - 5 mg)

A.F.S. Germany
Fax 49 - 7071 - 73141

Syntenic LabTracker is the first of several software packages for molecular biologists to be released by Syntenic Limited over the next year. Syntenic LabTracker runs under Windows for the PC and consists of populated databases including restriction enzymes, methylases, modification enzymes and bacterial strains. Syntenic LabTracker therefore provides a focal point to which information can be constantly added to and taken from, allowing information and ideas to be readily shared among the different members of a lab. A 'Commercial Sources' database allows companies which sell the product(s) of interest to be easily identified. Syntenic LabTracker will be available from Syntenic Limited during April 1995. Price £100 + VAT. For more information or if your company would like to be included in Syntenic LabTracker please contact us directly without delay. Syntenic Limited, Suite 2, Dolphin House, Sutton Harbour, Plymouth, PL4 0DW, UK. Tel/Fax: 01752 256 473 (UK) +44 1752 256 473 (Rest of world) email: syntenic@trump.demon.co.uk

DNA SEQUENCING

Using an ABI 373
Automated Sequencing
System

£45 per run

- PCR products
- M13 inserts
- plasmid inserts
- single stranded DNA
- double stranded DNA
- 300-400 bases of sequence

Please contact us for more information

Molecular Medicine Unit

King's College School of Medicine & Dentistry
Tel: 071- 346 3126 Fax: 071- 733 3877

LIPOSOMES

REPRODUCIBLE — HOMOGENEOUS —
EXTREMELY RAPID



Our new LIPOSOMAT guarantees rapid (5ml within 60 minutes) preparation of uniformly sized unilamellar liposomes. Liposome size is selectable between 25 and ca. 600 nm diameter. Sample volume is normally 5-10ml or up to 200ml if used in combination with an additional instrument. Lipid concentration can be up to 300 mg/ml.

DIANORM

P.O. Box 650126, D-81215 Munich
For USA call 1-800 Liposome